

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I



TESIS DOCTORAL

**INTERACCIONES PROTEÍNA-MEMBRANA
UNA APROXIMACIÓN ESTRUCTURAL POR RMN UTILIZANDO SISTEMAS
MODELO**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR



Soraya Serrano Serrano

Directoras

Marta Bruix Bayés
María Ángeles Jiménez López

Madrid, 2015

INTERACCIONES PROTEÍNA-MEMBRANA: UNA APROXIMACIÓN ESTRUCTURAL POR RMN UTILIZANDO SISTEMAS MODELO

TESIS DOCTORAL DE 
Soraya Serrano Serrano 
Madrid 2014

UNIVERSIDAD COMPLUTENSE DE MADRID

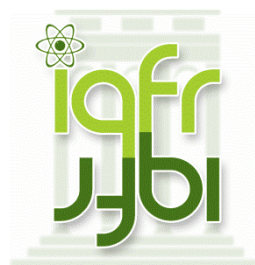
FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Bioquímica y Biología Molecular I



CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



**INTERACCIONES PROTEÍNA-MEMBRANA:
UNA APROXIMACIÓN ESTRUCTURAL POR
RMN UTILIZANDO SISTEMAS MODELO**

Soraya Serrano Serrano

MADRID 2014

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Bioquímica y Biología Molecular I

**INTERACCIONES PROTEÍNA-MEMBRANA:
UNA APROXIMACIÓN ESTRUCTURAL POR
RMN UTILIZANDO SISTEMAS MODELO**

Memoria presentada por

Soraya Serrano Serrano

Directoras de Tesis

Dra. Marta Bruix Bayés

Dra. María Ángeles Jiménez López

Tutora

Dra. Rosalía Rodríguez García

El trabajo ha sido realizado en el Grupo de Estructura, Dinámica e Interacciones de Proteínas por RMN, en el Departamento de Química Física Biológica del Instituto de Química Física "Rocasolano" (CSIC), bajo la ayuda Predoctoral de Formación de Personal Investigador de referencia BES-2009-018112 asociada al proyecto de investigación CTQ2008-00080.

A mis padres

A Cayetano

A Juana

A Sey

***“La verdadera ciencia enseña,
por encima de todo, a dudar y a ser ignorante”.***

Miguel de Unamuno (1864-1936)

AGRADECIMIENTOS

El trabajo que aquí se presenta se ha llevado a cabo en el grupo de Estructura, Dinámica e Interacciones de Proteínas por RMN, en el Departamento de Química Física Biológica del Instituto de Química Física “Rocasolano”, en el Consejo Superior de Investigaciones Científicas, bajo la dirección de la Prof. Marta Bruix Bayés y la Dra. M^a Ángeles Jiménez López a quienes quiero expresar mi más sincero y profundo agradecimiento por darme la oportunidad de estar hoy aquí, por enseñarme una ética de trabajo que es muy difícil de encontrar y esa pasión por la RMN que sólo ellas muestran. Muchas gracias por vuestros consejos, por vuestra paciencia y por vuestro constante apoyo.

Al Prof. Manuel Rico, el que ha sido y será el “jefe” de este grupo, por todo su esfuerzo y dedicación a todos nosotros.

Al Prof. Jorge Santoro por todo lo que nos ha enseñado sobre la RMN. Al Prof. Carlos González, por su gran apoyo, positividad y disposición. Al Dr. Douglas V. Laurents, por esa inquietud y pasión por la ciencia. Al Dr. José Manuel Pérez, por todos esos recursos y alternativas que guarda en el bolsillo. Al Prof. Alfredo Villasante, por habernos resuelto muchas dudas y por confiar tanto en nosotros, casi siempre, más que nosotros mismos.

A Cristina López y al Dr. Miguel Treviño, agradecerles la gran ayuda que me han prestado en diferentes momentos de este periodo. A Luis de la Vega, por su gratificante y entretenida compañía.

De este grupo de profesionales, solo podía salir un grupo de gente buena que tanto nos ha facilitado la vida a los estudiantes. Al Dr. David Pantoja Uceda, por siempre encargarse de todo, por tener la respuesta o el scrip que puede ahorrarte 5 horas de trabajo o enseñarte que existe la posibilidad de hacerlo, el mejor compañero que se puede tener. A la Dra. Irene Gómez Pinto, por su buen humor y su ánimo, transmitiéndote una energía que parece no tener fin. A la Dra. M^a Flor García Mayoral, a la que he acibillado a preguntas y tanto me ha ayudado y de la que se pueden aprender muchas cosas, pero sobre todo la dedicación.

Mi agradecimiento más tierno es para mis compañeros en esta andadura. A los del principio: Inés, Mar, Jorge Pedro, Fernando. A la gente de “arriba”: Cris, Lupe, Noemi, Laura, Silvia, Rosa, Javi, Santi.... Sin duda, una de las mejores cosas que me ha proporcionado esta Tesis es haber conocido a “los Rocalanis”, merecedores del título de TITANES, gente a la que admiro como científicos, pero todavía más como personas. A Palma, tanto corazón y trabajo era difícil que se juntasen en una persona y ahí está. A Santi, por ser tan crack y por su compañía. A Nerea, por cada momento compartido y por lo que se agradece tener a alguien al lado que tiene las cosas claras. A Didi, somos de la familia, otra intensa, entonces ¿qué voy a decir de ella?, ¡gracias por tu humor, tu inteligencia, por estar!. A Yasmina, todo el mundo debería tener una “Yasmina” en su vida, en su trabajo, en su labo... siempre está ahí apoyándote y,

seguramente sin darse cuenta, decirte lo que necesitas oír, jeso sí es un don!. A Miguel A. Pardo, que es todo corazón. A Miguel Mompeán, por cada momento de relax “café y piti” y por ser del selecto club “la cofradía de la última hora”. A Mon, que es un fenómeno, con su ingenio e inteligencia, por supuesto por su compañía en los “coffee break”. A Héctor, que nos recuerda que aquí todavía hay mucho que decir. Por último, a Mr. Miguel Garavís, el Rey de la gente de “buen vivir”, por ser el mejor en muchas muchas cosas pero sobre todo por ayudarme, animarme, distraerme, acompañarme, hacerme reír y sobre todo por haber estado aquí en esta etapa de mi vida; ¿qué voy yo a hacer ahora?, como te voy a echar de menos Miguel, a ti y a los otros 12.

I am very grateful to Professors Isabella C. Felli and Roberta Pierattelli for all the care, passion and patience they showed me during my stay at CERM. I wish to express my gratitude also for having given me the possibility to learn the fundamentals of research from their group and for their many insightful comments and suggestions. I would like to thank all people and friends in this center (Marie-Emmanuelle, Serena, Dani, Gianlu, Carmelo, Chiara, Masiek, Carolina, Angelo, Tomas.....) for having welcomed me, for the kindness and willingness they showed me, and for the funny times we spent together during toughest days. A special thank goes to Alessandro Piai, quanto mi manchi Ale!! la persona piú intelligente che mai ho conosciuto, ma tifoso della Juve!!!. Anche il piú grande ringrazio per la gente da fuori dalla ricerca: Sara, Anna e Ivan, Agnesca, Triziana ... grazie per avermi accettato nella vostra grande familia, soprattutto a Ale e Andre, siete così grandi ragazzi!!!.

A mi gente de la Universidad Autónoma de Madrid. A Mery, por su ilusión y ganas a la hora de empezar algo nuevo que resultan contagiosas. A Carmen y Laurich, por su buen humor. A Halbertoak, por su aptitud en la vida divertida, particular y desorientada. A Paco, porque es pura pasión y si la siente, nunca se le resistirá nada. A Edu, del cual uno aprende de verdad lo que es la responsabilidad y el trabajo duro. A Emilio, el más grande, el más equilibrado (¡de mayor quiero ser como él!). A Loren, no hay un solo día que se pase con él y no se aprenda algo, aunque no sean matemáticas. A Víctor, al que tanto admiro y tanto echo de menos, mi mentor ¡y no sólo en ciencia!. A Antón, el peso pesado, la persona a la que le confiaría todo y tengo razones más que de sobra para decirlo, porque me lo ha demostrado (excepto mi coche o el de mi padre ☺), te quiero “Kari”. A Natxo, mi amigo, mi hermano, el hombro en el que apoyarme, parte de mi alma y sobre todo mi mitad corpúsculo, gracias grandullón, ¡¡¡sin ti estaría perdida!!!.

A toda la gente del “Phantoms Team” o cercanos. A José Luis Roldan, a Joaquín, a Viviana, a Juan José Sáenz, a Perico, a Luis Froufe, a Manuel Marqués, a Antonio García, a José Costa ... sobre todo a mi Carmenchu y a Maite, por su ánimo y apoyo para embarcarme en esto, a pesar de ser “mayorcita”.

A todos mis amigos de la sierra, sobre todo de Valdemanco, mi casa. A Vicky y Sara; Javi, Héctor, Sergio, Rober, Guille, Iván, a María; a Patry y Arturo, por recordarme que otra vida es posible; a mi buen e incondicional Buji, por ser tan liante y haber estado siempre ahí y a Maica; a Dani, por ser como es de especial y a Laura; a Juampi y Nines; a Sari, por su confianza; a Edu y Paloma; a Oscar, por lo que cuida a Sey y porque siempre tiene un “claro, cuando te venga bien” cuando le pido algo y ¡¡son muchas veces!!; al bueno de Jorge, alias “Varita” y a Marga, vaya par!!; a Fran y a tantos otros ... pero sobre todo a Alberto, esa persona que cambió mi vida para siempre y no hay palabras que expresen la gratitud y el amor que le tengo; gracias por tus consejos, por tú apoyo y tú paciencia incondicional.

A mi familia, que son muchos y muy buenos. A mis tías/tíos, primas/primos, a los/las peques.... Tengo que mencionar a mi tía Neli, “mi primera madre”, por estar siempre para todo y por haberme enseñado tantísimas cosas. A mis abuelos, ¡los mejores!, cariñosos y humildes, que siempre estuvieron, y están, orgullosos de su nieta.

A Kike, por cada momento compartido, cada palabra, cada consejo, cada crítica... cada canción. “La música de mi vida”. Simplemente, gracias por estar.

A mis padres, gracias por haberme enseñado el valor del esfuerzo, la superación y el respeto. Gracias por ser como sois de perseverantes, trabajadores, cabezones, temperamentales y un montón de cosas más, pero sobre todo cariñosos y valientes. Cuando uno tiene unos padres como vosotros se atreve con cualquier cosa porque sabe que siempre, siempre, siempre estarán ahí para ayudarte, también para regañarte, pero sobre todo para animarte y jamás dejarte pasar ni frío ni hambre, ¿verdad papá?. Mamá, gracias por recordarme cada día lo que me quieres. Os quiero tanto que no lo puedo expresar, pero no sabéis la suerte que tenemos Sey y yo (y ahora Yako, que a pesar de ser un perro es el “niño bonito” de la casa). Gracias por hacerme sentir la persona más afortunada del mundo. Este trabajo es tan vuestro como mío.

Gracias a mi pequeña, mi princesa, mi vida y mi alma, a veces mi desesperación también. Posiblemente la única persona que me conoce de verdad y aun así me sigue queriendo, ¿cómo se puede ser tan buena?. El universo tiene suerte de que haya gente como ella porque compensa a unos 20 como yo y le devuelve a uno la fe en el ser humano. Gracias Sey, ¡¡te adoro!! Pero sobre todo ¡¡gracias por traerme de vuelta cuando me despisto!!.

Gracias a muchas otras personas que he ido encontrando en el camino, con las que he vivido grandes momentos y me han enseñado y ayudado muchísimo, demasiadas para citar aquí, ya me he pasado..... ¡Qué suerte tengo!.

Soraya

“It’s impossible”, said pride.

“It’s risky”, said experience.

“It’s pointless”, said reason.

“Give it a try”, whispered the heart.

ÍNDICE

ABREVIATURAS

ABSTRACT

CAPÍTULO I: INTRODUCCIÓN GENERAL.....	1
1. IMPORTANCIA DE LAS INTERACCIONES EN BIOLOGÍA.....	4
2. TIPOS DE INTERACCIONES.....	5
2.1. Proteína-proteína.....	5
2.2. Proteína-ácido nucleico.....	6
2.3. Proteína-carbohidrato.....	7
2.4. Proteína-lípido.....	8
3. MEMBRANAS BIOLÓGICAS.....	10
3.1. Composición.....	11
3.1.1. Lípidos.....	11
3.1.2. Proteínas de membrana.....	12
3.1.3. Carbohidratos.....	12
3.2. Características relevantes de las proteínas de membrana.....	13
4. IMPORTANCIA DE LAS INTERACCIONES PROTEÍNA-LÍPIDO Y PROTEÍNA-MEMBRANA.....	14
5. IMPORTANCIA DE LAS INTERACCIONES PROTEÍNA-PROTEÍNA. PROTEÍNAS INTRÍNECAMENTE DESORDENADAS.....	15
5.1. Paradigma estructura-función.....	15
5.2. Proteínas intrínsecamente desordenadas.....	16
6. MÉTODOS EXPERIMENTALES PARA EL ESTUDIO DE LAS INTERACCIONES	17
7. LA RMN COMO MÉTODO PARA ESTUDIAR INTERACCIONES, FUERTES Y TRANSITORIAS.....	18
8. PROBLEMÁTICA Y SOLUCIONES. MIMÉTICOS TANTO DE MEMBRANAS COMO DE PROTEÍNAS.....	21
9. OBJETIVO GENERAL.....	24
10. BIBLIOGRAFÍA.....	26

CAPÍTULO II: DESARROLLO DE LA SECUENCIA INTER-CAN, DE DETECCIÓN DE ¹³C, PARA EL ESTUDIO DE ESTADOS DESPLEGADOS.....	29
---	-----------

1. INTRODUCCIÓN.....	31
1.1. Problemas de la detección de ¹ H y alternativas.....	32
1.2. Detección directa de ¹³ C.....	33
1.2.1. Característica de la detección directa de ¹³ C.....	34
1.2.1.1. Sensibilidad.....	34
1.2.1.2. Resolución y anchura de línea.....	35

1.2.1.3.	Acoplamiento homonuclear.....	36
1.2.2.	Aplicaciones de la detección directa de ^{13}C para biomoléculas.....	38
1.2.2.1.	Proteínas paramagnéticas.....	40
1.2.2.2.	Sistemas de gran tamaño y eficiente relajación de espín.....	40
1.2.2.3.	Proteínas y regiones desestructuradas.....	41
1.2.2.4.	Sistemas con intercambio rápido con el solvente.....	42
1.3.	Adquisición de datos en RMN.....	43
1.4.	Objetivo.....	45
2.	MATERIALES Y MÉTODOS.....	46
2.1.	Cálculo de los desplazamientos químicos del random coil.....	46
2.2.	Muestra de RMN empleada.....	46
2.3.	Experimentos de RMN.....	46
2.4.	Análisis y asignación espectral.....	47
3.	RESULTADOS.....	48
3.1.	Limitación en la detección directa de ^1H . Necesidad de otras aproximaciones.....	48
3.2.	Exploración de las zonas espectrales de interés en detección de ^{13}C para proteínas sin estructura.....	50
3.3.	Mejoras en la dispersión de señales con la detección de ^{13}C y NUS.....	53
3.4.	Experimento de RMN inter-CAN.....	57
4.	DISCUSIÓN.....	61
4.1.	Nuevas propuestas para el estudio por RMN de IDPs. La secuencia inter-CAN.....	62
4.2.	Aplicación a la α -sinucleína.....	64
5.	CONCLUSIONES.....	67
6.	BIBLIOGRAFÍA.....	68
CAPÍTULO III: PÉPTIDOS BOOMERANG.....		73
1.	INTRODUCCIÓN.....	75
1.1.	Los glicosaminoglicanos.....	75
1.2.	Reconocimiento proteína/heparina.....	76
1.3.	Lipopolisacáridos.....	79
1.4.	Los péptidos antimicrobianos β -boomerang adoptan el motivo de unión a heparina clip CPC' en micelas de lipopolisacáridos (LPS).....	80
1.5.	Objetivo.....	84
2.	MATERIALES Y MÉTODOS.....	85
2.1.	Síntesis de los péptidos.....	85
2.2.	Requisitos de la muestra.....	85

2.3.	Preparación de las muestras.....	86
2.3.1.	Péptidos en solución acuosa.....	86
2.3.2.	Titulaciones con heparina.....	86
2.3.3.	Titulaciones con Fondaparinux.....	86
2.3.4.	Péptidos en presencia del cosolvente TFE.....	87
2.4.	Adquisición de los espectros de RMN.....	87
2.5.	Procesado de los espectros de RMN.....	87
2.6.	Asignación espectral.....	88
2.7.	Determinación de la estructura secundaria.....	89
2.8.	Análisis de los resultados.....	90
3.	RESULTADOS.....	91
3.1.	Diseño de los péptidos.....	91
3.2.	Estructura secundaria de los péptidos en disolución acuosa y en presencia de TFE.....	92
3.3.	Titulación de YI12 con heparina.....	97
3.4.	Interacciones del péptido YI12 con Fondaparinux.....	99
3.5.	Interacciones de los péptidos con residuos catiónicos centrales mutados con Fondaparinux.....	101
3.6.	Interacciones de los péptidos mutados en los extremos con Fondaparinux...	106
3.7.	Interacciones del péptido YI12M9 con Fondaparinux.....	109
4.	DISCUSIÓN.....	111
4.1.	Estructura de los péptidos en agua y en la mezcla agua/TFE.....	111
4.2.	Interacciones entre los péptidos y Fondaparinux en presencia de TFE.....	113
4.3.	Conservación del motivo grapa CPC'.....	115
5.	CONCLUSIONES.....	116
6.	BIBLIOGRAFÍA.....	118

CAPÍTULO IV: ESTRUCTURA DE PÉPTIDOS QUE MIMETIZAN DISTINTAS REGIONES DEL ECTODOMINIO DE LA GLICOPROTEÍNA GP41 DEL VIRUS VIH-1..... 121

1.	INTRODUCCIÓN.....	123
1.1.	El Virus de Inmunodeficiencia Humana (VIH).....	123
1.2.	El complejo Env: diana de anticuerpos frente a VIH.....	125
1.3.	Relevancia de gp41 en el mecanismo de fusión de membranas.....	126
1.4.	Estructura de gp41 y de sus dominios.....	127
1.5.	La región FP del ectodominio de gp41.....	128
1.6.	La región MPER del ectodominio de gp41.....	130
1.7.	La región TMD de gp41.....	133

1.8.	Objetivo.....	134
2.	MATERIALES Y MÉTODOS.....	135
2.1.	Preparación de las muestras.....	135
2.2.	Adquisición de los espectros de RMN.....	136
2.3.	Procesado de los espectros de RMN.....	136
2.4.	Asignación espectral.....	136
2.5.	Determinación de la estructura secundaria.....	136
2.6.	Cálculo de estructuras tridimensionales.....	137
2.7.	Visualización y análisis de las estructuras.....	137
3.	RESULTADOS.....	138
3.1.	Péptidos quimera de las regiones FP y MPER: HybK3 y scrHybK3.....	138
3.1.1.	Diseño de los péptidos.....	138
3.1.2.	Espectros de RMN.....	140
3.1.3.	Análisis cualitativo de la estructura secundaria.....	140
3.1.4.	Estructuras tridimensionales de los péptidos.....	142
3.2.	Péptidos de las regiones MPER y TMD.....	149
3.2.1.	Diseño de los péptidos MPERp, CpreTM y TMDp.....	149
3.2.2.	Espectros de RMN.....	153
3.2.3.	Análisis cualitativo de la estructura secundaria.....	153
3.2.4.	Estructuras tridimensionales de los péptidos.....	157
4.	DISCUSIÓN.....	162
4.1.	Estructura de la región FP.....	162
4.2.	Estructura de región MPER e implicaciones para la inmunogenicidad.....	165
4.3.	Interacción entre las regiones FP y MPER.....	168
4.4.	Estructura de la región TMD.....	170
4.5.	Implicaciones de la estructura de los péptidos en el mecanismo de fusión de las membranas.....	171
5.	CONCLUSIONES.....	177
6.	BIBLIOGRAFÍA.....	178

CAPÍTULO V: ESTRUCTURA TRIDIMENSIONAL DEL ZIMÓGENO FLG DE ONCONASA® ACTIVABLE POR LA PROTEASA DEL VIH-1..... 183

1.	INTRODUCCIÓN.....	185
1.1.	Ribonucleasas y Onconasa®.....	185
1.2.	Zimógenos de RNasas.....	187
1.3.	El zimógeno de la Onconasa®.....	189
1.4.	Diseño y optimización del zimógeno ONCFLG de la Onconasa®.....	190

1.5	Objetivo.....	193
2.	MATERIALES Y MÉTODOS.....	195
2.1.	Preparación de la muestra para RMN del zimógeno ONCFLG.....	195
2.2.	Hidrólisis del zimógeno ONCFLG.....	195
2.3.	Adquisición de los espectros de RMN para el cálculo de estructura.....	197
2.4.	Procesado y análisis de los espectros de RMN.....	197
2.5.	Asignación espectral.....	198
2.6.	Cálculo de estructuras.....	199
2.7.	Análisis de estructuras.....	199
2.8.	Minimización.....	200
2.9.	Validación y depósito.....	200
2.10.	Interacción de ONCFLG con el análogo de sustrato d[UGG] ₃	200
2.11.	Construcción del modelo estructural del zimógeno hidrolizado.....	201
2.12.	Adquisición de los espectros de RMN para la dinámica.....	201
2.13.	Adquisición de los espectros de RMN para el estudio de estabilidad conformacional.....	202
3.	RESULTADOS.....	204
3.1.	Caracterización preliminar de la proteína.....	204
3.2.	Asignación espectral.....	205
3.3.	Estructura del zimógeno ONCFLG.....	207
3.3.1.	Estructura secundaria.....	207
3.3.2.	Estructura tridimensional del zimógeno ONCFLG.....	210
3.3.3.	Análisis de la estructura en relación a la actividad ribonucleolítica....	219
3.4.	Unión de ONCFLG a un análogo de sustrato.....	222
3.5.	Hidrólisis del zimógeno ONCFLG.....	223
3.6.	Estructura del zimógeno ONCFLG hidrolizado.....	224
3.7.	Dinámica del zimógeno hidrolizado y no hidrolizado.....	228
3.8.	Estabilidad conformacional.....	229
4.	DISCUSIÓN.....	232
4.1.	Estructura de la secuencia insertada en el zimógeno ONCFLG.....	233
4.2.	Bases estructurales de la actividad del zimógeno ONCFLG.....	234
4.3.	Bases estructurales de la actividad del zimógeno ONCFLG hidrolizado.....	236
5.	CONCLUSIONES.....	228
6.	BIBLIOGRAFÍA.....	240
	CAPÍTULO VI: RESUMEN Y CONCLUSIONES GENERALES.....	245

APÉNDICE.....	251
FIGURA A.1. $\Delta\delta C_\alpha$ y $\Delta\delta C_\beta$ de HybK3 y scrHybK3 en 25% HFIP.....	253
FIGURA A.2. $\Delta\delta C_\alpha$ y $\Delta\delta C_\beta$ de CpreTM en 25% HFIP.....	254
FIGURA A.3. $\Delta\delta H_\alpha$ de HybK3, MPER y CpreTM en 20 mM DPC.....	254
TABLA A.1. Desplazamientos químicos de HybK3 en 25% HFIP.....	255
TABLA A.2. Desplazamientos químicos de scrHybK3 en 25% HFIP.....	257
TABLA A.3. Desplazamientos químicos de HybK3 en 20 mM DPC.....	259
TABLA A.4. Desplazamientos químicos de scrHybK3 en 20 mM de DPC.....	261
TABLA A.5. Desplazamientos químicos de MPERp en 20 mM DPC.....	263
TABLA A.6. Desplazamientos químicos de MPERp en 25% HFIP.....	265
TABLA A.7. Desplazamientos químicos de CpreTM en 20 mM DPC.....	267
TABLA A.8. Desplazamientos químicos de CpreTM en 25% HFIP.....	269
TABLA A.9. Desplazamientos químicos de TMD en 25% HFIP.....	271
TABLA A.10. Desplazamientos químicos del zimógeno ONCFLG.....	273

ABREVIATURAS

1D	Monodimensional
2D	Bidimensional
3D	Tridimensional
4D	Tetradimensional
6-HB	<i>Six-Helix Bundle</i> (manejo de 6 hélices)
Å	Ángstrom
Ahx	Ácido 6-amino hexanoico
AP	<i>Anti-phase</i> (anti-fase)
BMRB	<i>Biological Magnetic Resonance Bank</i>
bNAbs	<i>Broadly Neutralizing Antibodies</i> (anticuerpos monoclonales neutralizantes)
C'	Carbono del grupo carbonilo del enlace peptídico
CD	<i>Circular Dichroism</i> (dicroísmo circular)
CHR	<i>C-terminal Heptad Repeat</i>
COSY	<i>COrelated SpectroscopY</i>
CPC'	Motivo catiónico-polar-catiónico
CSM	<i>Chemical Shift Mapping</i> (mapeo de perturbación de los desplazamientos químicos)
CSP	<i>Chemical Shift Perturbation</i> (cartografiado de desplazamientos químicos)
CT	<i>Cytoplasmic Tail</i> (tallo citoplasmático)
Da	Dalton
DIPAP	Doble IPAP
DPC	Dodecilmfosfocolina
DMSO	Dimetil sulfóxido
DSC	<i>Differential Scanning Calorimetry</i> (calorimetría diferencial de barrido)
DSS	2,2-dimetil-2-silapentano-5-sulfonato sódico
EDTA	Ácido etilendiaminotetraacético
FID	<i>Free Induction Decay</i>
FP	<i>Fusion peptide</i> (péptido de fusión)
FTIR	<i>Fourier Transform Infrared Spectroscopy</i> (espectroscopía infrarroja por transformada de Fourier)
FX	Fondaparinux
GAG	Glicosaminoglicanos
gp	Glicoproteína
HAART	<i>Highly Active Antiretroviral Therapy</i> (tratamiento antirretroviral de gran actividad)
HEPES	Ácido 2-[4-(2-hidroxietil)-1-piperacil] etanosulfónico
HFIP	1,1,1,3,3,3-hexafluoro-isopropanol
HSQC	<i>Heteronuclear Single Quantum Coherence spectroscopy</i>
Hz	Hercios
IDPs	<i>Intrinsically Disordered Proteins</i> (proteínas intrínsecamente desestructuradas)
IDRs	<i>Intrinsically Disordered Regions</i> (regiones intrínsecamente desestructuradas)
IP	<i>In-phase</i> (en-fase)
IPAP	<i>In- phase anti-phase</i> (en-fase anti-fase)

$^nJ_{AB}$	Constante de acoplamiento a n enlaces entre los átomos A y B
K_d	Constante de equilibrio de disociación
k_{off}	Constante de velocidad de la reacción de disociación de un complejo
k_{on}	Constante de velocidad de la reacción de formación de un complejo
IPAP	<i>In-phase Anti-phase</i> (en-fase anti-fase)
LPS	Lipopolisacáridos
MA	Matriz
MA/CA	Matriz/Cápside
MPER	<i>Membrane Proximal External Region</i> (región externa próxima a la membrana)
MFT	<i>Multidimensional Fourier Transform</i> (transformada de Fourier Multidimensional)
NHR	<i>N-terminal Heptad Repeat</i>
NOE	<i>Nuclear Overhauser Effect</i> (efecto Overhauser nuclear)
NOESY	<i>Nuclear Overhauser Effect Spectroscopy</i>
NUS	<i>Non-uniform sampling</i> (muestreo no-uniforme)
ONC	Onconasa® nativa
ONCFLG	Zimógeno FLG
PDB	<i>Protein Data Bank</i>
PFG	<i>Pulsed Field Gradients</i> (pulsos de gradiente de campo)
ppm	partes por millón
RMN	Resonancia Magnética Nuclear
RDC	<i>Residual Dipolar Coupling</i> (constante dipolar residual)
RMSD	<i>Root Mean Square Deviation</i> (desviación cuadrática media)
RNasa A	Ribonucleasa A
SDS	Dodecilsulfato sódico
SIDA	Síndrome de Inmunodeficiencia Adquirido
SMFT	<i>Sparse Multidimensional Fourier Transform</i>
TFE	2,2,2-trifluoroetanol
TMD	<i>Transmembrane domain</i> (dominio transmembrana)
TOCSY	<i>TOTAL Correlated Spectroscopy</i>
TROSY	<i>Transverse Relaxation-Optimized Spectroscopy</i>
VIH	Virus de Inmunodeficiencia Humana
δ	Desplazamiento químico
δ_{rc}	Desplazamiento químico random coil

ABSTRACT

Nuclear Magnetic Resonance (NMR) spectroscopy has applications in multiple areas of science, especially in chemistry, biochemistry and biomedicine. NMR is a very versatile technique for determining the structure, dynamics and interactions of organic molecules and biomolecules in solution. Also, NMR is a technique experimenting rapid and continuous advances, both by instrument technology and by the development of new experimental techniques.

In this context, chapter II of this Thesis addresses the development of a new pulse sequence, which provides a new tool for the study of intrinsically disordered proteins (IDPs). This is of great relevance because of the complex conformational behavior of IDPs, and the increasing number of known IDPs.

Meanwhile, chapters III-V exploit the NMR versatility for the characterization and structural determination of biomolecules and their interactions. Structural characterization and dynamics of these interactions with atomic resolution is essential to understand their functionality within the cell. This information can be of practical use, for instance, in the development of new drugs and new therapeutic strategies for the treatment of various diseases. Within this broad field, chapters III-IV aim to characterize the interactions of two particular biological systems using both protein and membrane mimetics. The two systems studied include boomerang peptides as heparin-binding models, in chapter III, and peptides derived from the gp41 protein in presence of simple mimetic of membrane environments, an apolar medium and micelles, in chapter IV. In chapter V, a complete characterization of a circularly permuted onconasa variant, ONCFLG zymogen, is performed. The internal dynamics and the conformational stability at the residue level are determined. The findings on these systems has provided us additional information to that available before starting this work, which allow us to derive structure-function implications.

In brief, the work done in this Thesis has provided us a new experimental NMR method applicable to intrinsically disordered protein, an emerging area in the protein field, and new data about three biological systems, which has improved our understanding of the molecular basis of their biological functions.

Capítulo I

Introducción general

Capítulo I:

Introducción

General

Los procesos vitales de las células se rigen por las interacciones entre biomacromoléculas particulares. De hecho, para poder ejercer su función biológica la mayoría de las proteínas requieren interactuar con otras especies moleculares, que pueden ser otras macromoléculas o ligandos pequeños. Para poder comprender las bases de la vida a nivel molecular, el primer paso es conocer la estructura de las moléculas individuales y posteriormente estudiar las reacciones entre las macromoléculas biológicas y sus compañeros de interacción. La caracterización estructural y dinámica a resolución atómica de estas interacciones nos permitirá entender los procesos celulares en los que intervienen y nos proporcionará información que puede concretarse, por ejemplo, en el desarrollo de nuevos fármacos y nuevas estrategias terapéuticas para el tratamiento de diversas enfermedades.

Dado el gran potencial de la espectroscopía de Resonancia Magnética Nuclear, RMN, para la caracterización y determinación estructural y dinámica a resolución atómica de macromoléculas, en el trabajo que aquí se presenta, se ha utilizado como herramienta para estudiar distintos sistemas biológicos a partir de proteínas concretas y sus interacciones con miméticos de membranas. Los resultados obtenidos aportan información complementaria a la disponible actualmente y resultan indispensables para comprender los procesos biológicos en los que intervienen los péptidos y proteínas objeto de este estudio.

1. IMPORTANCIA DE LAS INTERACCIONES EN BIOLOGÍA

Las interacciones entre biomacromoléculas representan el lenguaje de los sistemas biológicos de manera que a través de ellas, las macromoléculas se comunican y ejercen su función. La mayoría de las interacciones celulares involucran y están dirigidas por proteínas que son las principales responsables de cualquier proceso biológico dentro de los organismos (Cosic 1994).

Además de los enlaces covalentes, que constituyen el andamiaje de proteínas, ácidos nucleicos, polisacáridos, lípidos, etc., existe un tipo adicional de interacciones interatómicas. Estas interacciones, denominadas débiles, son mucho menos energéticas y por lo tanto pueden romperse o formarse con mayor facilidad. Las interacciones débiles juegan un papel biológico esencial. Estas interacciones pueden ser de varios tipos: enlaces de hidrógeno, interacciones electrostáticas, interacciones apolares o hidrófobas y fuerzas de van der Waals. Muchos de los acontecimientos a nivel molecular que tienen lugar en las células vivas, entre los que por su importancia biológica podríamos destacar la estructura tridimensional que adoptan las macromoléculas o la relación estereoespecífica entre la enzima y su sustrato, responden a este tipo de interacciones.

2. TIPOS DE INTERACCIONES

Dada la importancia de las proteínas en los procesos vitales de la célula, las interacciones en las que nos centraremos aquí son las interacciones débiles que se dan entre las proteínas y otros tipos de biomoléculas.

2.1. Proteína-proteína

Las interacciones proteína-proteína juegan un papel clave en la comprensión del complejo mundo de los procesos biológicos que se producen dentro de las células vivas, ya que las funciones celulares son el resultado de la acción coordinada de grupos de proteínas que interactúan entre sí. Procesos tan importantes e imprescindibles como la replicación del DNA o la señalización celular se llevan a cabo a través de las interacciones proteína-proteína.

En las interacciones proteína-proteína, el tipo de fuerzas que mantienen unidas distintas cadenas peptídicas son, en líneas generales, las mismas que estabilizan la estructura terciaria dentro de una misma proteína (interacciones hidrófobas, fuerzas de van der Waals, enlaces de hidrógeno y fuerzas iónicas entre las cadenas laterales de los aminoácidos).

Las cadenas participantes en las interacciones proteína-proteína pueden o no ser idénticas. Los oligómeros formados a partir de las interacciones entre proteínas idénticas dan lugar a homo-oligómeros. Estos complejos se pueden formar mediante la asociación a través de la misma superficie de los monómeros (como ocurre en el caso del represor Arc, código PDB 1B28) o bien pueden ensamblarse usando superficies distintas, sin simetría (al contrario de lo que ocurre en el caso anterior) pudiendo conducir a procesos de agregación inespecífica. Cuando las cadenas que se asocian son diferentes entre sí se forman los hetero-oligómeros (Monod et al. 1965).

Además de las distintas cadenas entre las que se dan las interacciones, se pueden distinguir dos tipos distintos de complejos atendiendo a si la formación del complejo está o no obligada en términos de función biológica. En una interacción obligatoria, la unidad estructural de una proteína oligomérica, protómero, no tiene estructura funcional por sí misma *in vivo*, como es el caso del represor Arc antes mencionado. Sin embargo, muchos de los complejos hetero-oligoméricos involucran interacciones no obligatorias ya que los protómeros existen independientemente de manera estable, como es el caso del complejo RhoA/RhoGAP (código PDB 1OW3).

En lo referente al tiempo durante el cual se establecen las interacciones, estas pueden darse durante tiempos largos o interaccionar temporalmente, durante un breve periodo que permita una modificación transitoria. Un ejemplo de interacción duradera proteína-proteína es el caso de barnasa, figura 1.1, una proteína pequeña que inhibe la actividad ribonucleolítica de la enzima barnasa a través de la formación de un complejo muy estable dentro de la célula que se mantiene hasta que la barnasa es secretada (Buckle et al. 1994). De manera contraria, en el complejo RhoA/RhoGAP, implicado en procesos de señalización (Graham et al. 2002), se dan interacciones transitorias proteína-proteína en el rango de los segundos (Perkins et al. 2010).

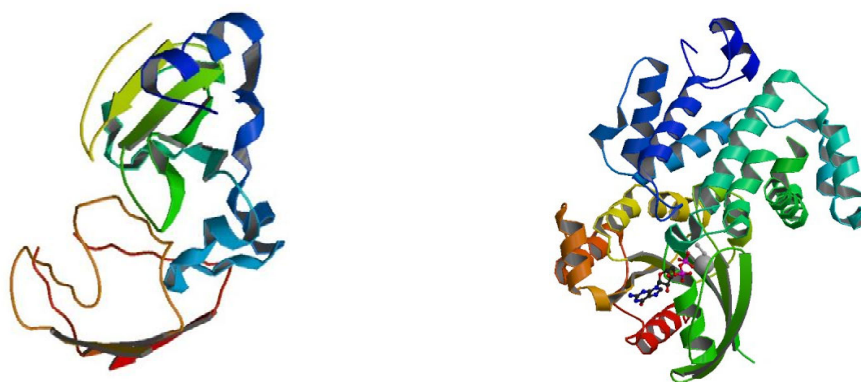


Figura 1.1. A la izquierda complejo barnasa-barstar, código pdb 1BRS y a la derecha complejo RhoA/RhoGAP con código pdb 1OW3.

2.2. Proteína-ácido nucleico

Algunas proteínas presentan dominios de unión a los ácidos nucleicos, con los que pueden interaccionar de forma específica o inespecífica. Muchos procesos biológicos importantes, como el transporte y la traducción del RNA, el empaquetamiento de DNA, la recombinación genética, replicación y reparación del DNA son controlados por la interacción de éstas dos clases de biomoléculas (Cai & Huang 2012).

Este tipo de interacciones se producen a través de las mismas fuerzas físicas que se han comentado en el caso de las interacciones proteína-proteína.

Un ejemplo importante de este tipo de interacción es el que se da con las enzimas que sintetizan o modifican el DNA, las cuales suelen establecer contactos extensos que, a menudo determinan que la molécula de DNA quede rodeada por la proteína en la región de interacción. Así por ejemplo, la aminoacil t-RNA sintetasa que se muestra en la figura 1.2 (Ruff et al. 1991), reconoce a un aminoácido concreto y al RNA de transferencia correspondiente y los une covalentemente.

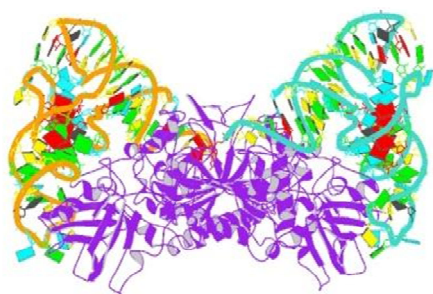


Figura 1.2. Estructura cristalina de la aspartil-tRNA sintetasa con la molécula de tRNA(Asp), código pdb 1ASY.

2.3. Proteína-carbohidrato

Los carbohidratos son una de las cuatro clases principales de biomoléculas y constituyen la mayor parte de la materia orgánica de la tierra, por lo que están implicados en multitud de funciones en los seres vivos. Normalmente se encuentran unidos a proteínas y lípidos, como ocurre con los carbohidratos de la superficie celular, involucrados de forma crucial en el proceso de reconocimiento intercelular. Estos procesos de reconocimiento molecular carbohidrato/proteína están relacionados con eventos tan relevantes como el plegamiento y maduración de las glicoproteínas, en la adhesión celular, en el control del crecimiento celular, etc.

Las interacciones que establecen son enlaces de hidrógeno, a menudo un grupo hidroxilo (OH) actúa simultáneamente de dador y de aceptor formando un enlace de hidrógeno cooperativo; interacciones de van der Waals, que se dan entre los átomos de las cadenas laterales de los residuos de la proteína con otros átomos del carbohidrato; interacciones entre regiones hidrófobas del carbohidrato y los residuos aromáticos. En ocasiones, los metales que están coordinados a la proteína interactúan con los carbohidratos.

Uno de los sistemas en los que encontramos este tipo de interacciones son las lectinas, proteínas responsables del reconocimiento de los carbohidratos que se encuentran en la superficie celular en bacterias, animales y plantas, figura 1.3, (Kolatkár & Weis 1996). Sólo parte de los grupos hidroxilo del carbohidrato forman enlaces de hidrógeno, lo que implica que la unión ocurra superficialmente. A pesar de que la afinidad de la proteína por cada monosacárido es baja, se consiguen afinidades altas al formarse complejos entre oligómeros de lectina y varios monosacáridos de la superficie celular.

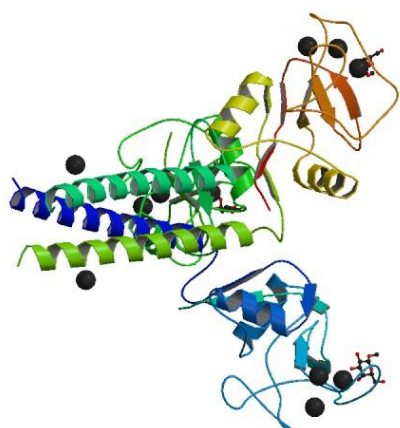


Figura 1.3. Representación de lecitina reconocida de forma específica por la galactosa o N-acetilgalactosamina; código pdb 1AFA. Las esferas representan átomos de Ca^{+2} .

2.4. Proteína-lípido

Los lípidos son las macromoléculas más abundantes en las membranas biológicas, por lo que las interacciones entre las proteínas y lípidos de membrana son fundamentales y están implicadas en multitud de procesos básicos como la respiración, la fotosíntesis, el transporte de solutos, la transducción de señales y la movilidad. La bicapa lipídica proporciona la matriz de las proteínas de membrana, por lo que se da una estrecha interacción lípido-proteína que permite mantener la barrera de difusión y conservar electroquímicamente sellada la membrana. Las moléculas de lípidos móviles pueden adherirse a la superficie de la proteína de membrana y ajustarse de forma flexible a los cambios conformacionales y reordenamientos estructurales que sufren estas últimas. Las superficies de los lípidos son fundamentalmente apolares, aunque con frecuencia presentan grupos cargados en un extremo. Sus interacciones con las proteínas se producen principalmente a través de interacciones hidrófobas y fuerzas de van der Waals. Los grupos polares pueden establecer otro tipo de interacciones como enlaces de hidrógeno e interacciones electrostáticas. Dentro de este tipo de interacciones encontramos numerosos ejemplos, como ocurre en la colesterol esterasa (Chen et al. 1998), que cataliza la hidrólisis específica de los ésteres de colesterol, figura 1.4.

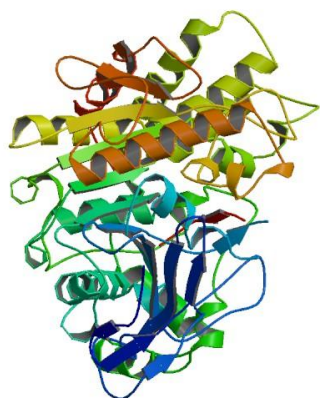


Figura 1.4. Estructura de la colesterol esterasa de páncreas bovino, código pdb 2BCE.

En base a las características comunes de los sitios de unión de lípidos de las estructuras disponibles de complejos lípido-proteína de membrana (Palsdottir & Hunte 2004), se han definido tres modos de unión. El primero, a través de la formación de una especie de carcasa de lípidos anulares unidos a la proteína de superficie, que se asemeja a la estructura de bicapas. En el segundo tipo, los lípidos de superficie no anulares se sumergen en las cavidades y hendiduras de las proteínas periféricas. Este tipo de interacción se observa con frecuencia en complejos formados por múltiples subunidades y ensamblajes multiméricos y están típicamente presentes en los sitios de contacto entre subunidades o monómeros adyacentes. Finalmente, hay interacciones entre lípidos y proteínas integrales que ocurren dentro de una proteína de membrana o en complejos de proteínas de membrana. Este tipo de interacción normalmente se da en las partes más periféricas de los lípidos como por ejemplo, con el grupo de cabeza por debajo del plano de la membrana y/o las cadenas de acilo no perpendiculares a la misma.

3. MEMBRANAS BIOLÓGICAS

Las membranas biológicas son superficies delgadas y flexibles, figura 1.5, que separan a las células y orgánulos del medio en el que se encuentran y mantienen la integridad territorial mediante la regulación de flujo de materia, información y energía entre el medio intra y extracelular (von Heijne & Rees 2008). Están compuestas, principalmente, por tres componentes: lípidos, proteínas y carbohidratos. Todas las membranas tienen una estructura general formada por dos capas de lípidos que constituyen una matriz en la que se incorporan proteínas de membrana (Singer & Nicolson 1972). Esta estructura es altamente fluida y la mayoría de lípidos y proteínas pueden moverse en el plano de la membrana. Estos lípidos y proteínas, se mantienen unidos entre si principalmente por interacciones no covalentes, mientras que los carbohidratos se unen a través de enlaces covalentes a algunos de estos lípidos y proteínas.

Existe una gran diversidad combinatoria resultado de la variabilidad de ácidos grasos (la combinación entre cadenas principales y grupos de cabeza diferentes, da lugar a miles de especies de lípidos según el tipo de célula (Han & Gross 2005)) y proteínas de membrana complementarias. En presencia de tantas especies moleculares distintas, las interacciones preferenciales entre ellas son inevitables (Engelman 2005).

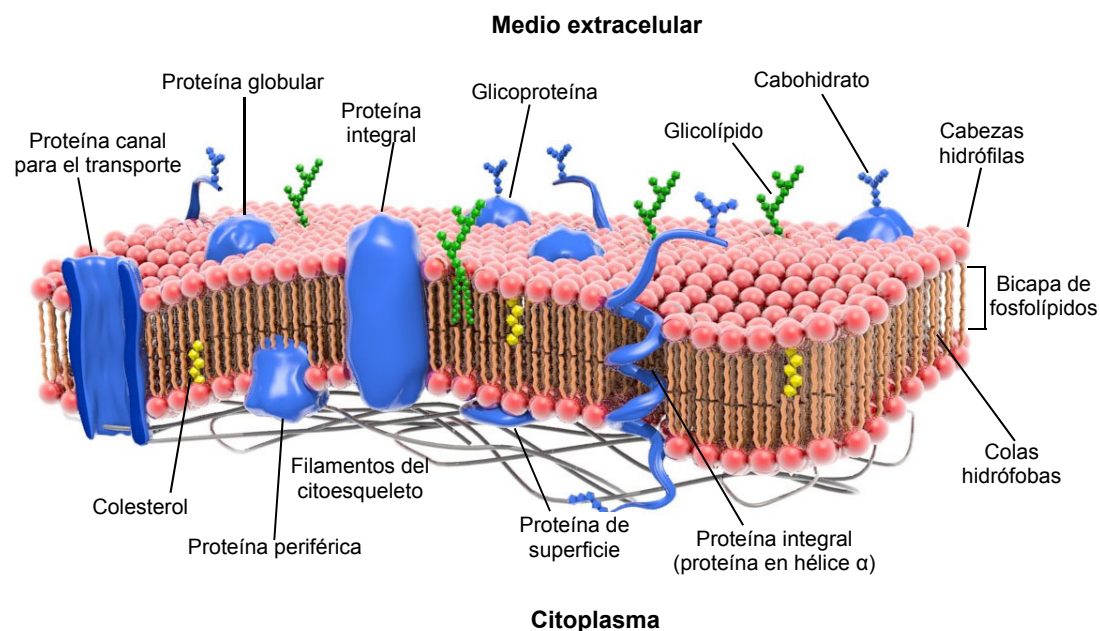


Figura 1.5. Representación de una membrana biológica en la que se puede observar la doble capa de fosfolípidos y los distintos tipos de proteínas y carbohidratos que la forman.

3.1. Composición

Cada especie, tejido, célula u orgánulo tiene una composición de membrana característica, tanto para lípidos como para las proteínas de membrana, que responde a las necesidades funcionales propias de esa membrana.

3.1.1. Lípidos.

Los lípidos son macromoléculas de gran importancia biológica, son insolubles en agua pero solubles en disolventes orgánicos como la acetona, el etanol, etc. En las membranas biológicas se encuentran tres tipos principales de lípidos: fosfolípidos, glicolípidos y colesterol, cada uno de los cuales desempeña distintos papeles en la membrana.

- Fosfolípidos. El tipo más común está formado por la molécula de glicerol que se une a dos cadenas de ácidos grasos, fosfato y colina. Las cadenas de ácidos grasos contienen entre 14 y 24 átomos de carbono. Normalmente, una de las cadenas es insaturada y contiene de uno a cuatro dobles enlaces en configuración *cis*.
- Glicolípidos. Son similares a los fosfolípidos y se diferencian de ellos en que contienen un azúcar, que puede ser glucosa o galactosa, en lugar de estar fosfatados. Los glicolípidos de las membranas celulares animales casi siempre contienen esfingosina, mientras que los de las membranas de bacterias y plantas principalmente contienen glicerol. En todos los casos, los glicolípidos se encuentran en la superficie externa de la membrana plasmática, con los carbohidratos expuestos hacia la superficie celular.
- Colesterol. Estructuralmente difiere bastante de los fosfolípidos y glicolípidos. El colesterol es un lípido esteroide que lleva una cadena alifática y un grupo hidroxilo. El colesterol se encuentra sólo en las membranas celulares de los mamíferos y en un tipo de microorganismo, el micoplasma. Es poco frecuente encontrarlo en membranas bacterianas o plantas.

3.1.2. Proteínas de membrana

Muchas de las funciones específicas de las membranas se llevan a cabo a través de proteínas. La cantidad y tipos de proteínas presentes en las membranas varían considerablemente de unas a otras. Cuanto más activo es el metabolismo de una célula, más proteínas encontramos en su membrana.

Hay varios tipos diferentes de proteínas asociadas a los lípidos en una membrana funcional:

- Proteínas integrales de membrana o transmembrana. Son aquellas proteínas que se extienden a lo largo de la bicapa lipídica. Las regiones hidrófobas de estas proteínas se encuentran embebidas en la bicapa, de manera que se dan interacciones con las colas hidrófobas de los fosfolípidos. Estas regiones, a menudo son helicoidales. Las regiones hidrófilas de estas proteínas se extienden a ambos lados de la membrana. Algunas proteínas transmembrana pueden estar unidas covalentemente a las cadenas de ácidos grasos que se encuentran insertados en la bicapa de fosfolípidos.
- Proteínas ancladas a la membrana. Son aquellas proteínas que no se extienden a lo largo de la membrana pero que están unidas covalentemente a las cadenas de ácidos grasos o de fosfolípidos, tanto del lado citosólico como del extracelular.
- Proteínas periféricas. En este grupo están incluidas aquellas proteínas que están unidas débilmente a la superficie interna o externa a través de interacciones no covalentes con otras proteínas de membrana.

3.1.3. Carbohidratos

Los carbohidratos unidos a proteínas o lípidos no suponen más del 10% del peso total de la membrana. La mayoría de las proteínas de membrana que están expuestas a la superficie celular están unidas covalentemente a carbohidratos. Estos carbohidratos a pesar de estar formados por cadenas relativamente cortas (entre 4 y 12 monómeros), presentan una gran variedad de estructuras. Cada tipo individual de célula muestra distintas características como consecuencia de los carbohidratos presentes en su membrana, que están involucrados, entre otros, en procesos de reconocimiento celular.

3.2. Características relevantes de las proteínas de membrana

Hay varias consideraciones que se deben tener en cuenta acerca de las propiedades estructurales de las proteínas de membrana:

- Las regiones transmembrana de las proteínas se caracterizan por ser ricas en aminoácidos hidrófobos.
- Dentro de los dominios transmembrana, las cadenas laterales que no están en contacto con los lípidos pueden interaccionar entre ellas en la interfase o pueden alinearse con un poro o cavidad. En este último caso los residuos a menudo son polares. Sin embargo, en las interfases proteína-proteína dentro de los dominios transmembrana, los residuos son típicamente apolares.
- La propensión a la formación de enlaces de hidrógeno de los grupos amida del esqueleto dentro de los segmentos transmembrana, normalmente están maximizados. Los segmentos transmembrana presentan casi siempre formas helicoidales o láminas beta (como parte de barriles).
- Las proteínas de membrana normalmente tienen dominios que están expuestos al agua. A veces estas regiones se pliegan automáticamente y otras forman giros que conectan segmentos transmembrana.
- La membrana está más organizada que un líquido isotrópico y es probable que su organización ayude a la estabilidad de las proteínas de membrana.
- El medio interfacial no es una ruptura aguda entre los dominios apolar y polar. Más bien, supone una transición gradual que se extiende unos 15 Å que separan la fase acuosa de una fase de hidrocarburos.
- Los dominios hidrófobos transmembrana de las proteínas de membrana no son estables cuando están expuestos al agua por lo que agregan si están en presencia de la misma.
- Aunque hay importantes excepciones (Marsh & Pali 2004; Gonen et al. 2005; Hunte 2005), la mayoría de interacciones lípido-proteína no son altamente específicas. Una proteína de membrana dada puede continuar siendo funcional cuando es reconstituida en otra bicapa lipídica de distinta composición. Esto es muy importante, ya que implica que una proteína de membrana puede conservar una estructura de tipo nativo bajo condiciones que se desvían considerablemente del medio de la membrana nativa, por lo que puede ser estudiada por RMN.
- Las proteínas de membrana normalmente rotan alrededor de la bicapa dentro de la membrana, con tiempos de correlación para la rotación en el orden de los microsegundos (Park et al. 2006).

4. IMPORTANCIA DE LAS INTERACCIONES PROTEÍNA-LÍPIDO Y PROTEÍNA-MEMBRANA

Las interacciones proteína-lípido son indispensables en una gran variedad de procesos biológicos, regulando actividades tan importantes para la célula como son la señalización celular y el paso de moléculas a través de la membrana celular (Kami et al. 2009).

No se dispone de mucha información acerca de los dominios de proteínas que interaccionan con la superficie de las membranas, debido a las dificultades prácticas que encontramos tanto en la expresión como en la purificación de las mismas, la baja solubilidad y la debilidad de las interacciones con los lípidos. Además de las dificultades metodológicas, encontramos una gran variedad de fosfolípidos y glicolípidos que pueden ser reconocidos específicamente o que influyen en las estructuras y dinámica de las bicapas y de las proteínas que interactúan con ellos. El hecho de que los mecanismos de reconocimiento de los lípidos puedan involucrar múltiples pasos, como la unión a la membrana e inserción, en los cuales las proteínas sufren cambios conformacionales, hace que para la caracterización de estas interacciones sea necesaria información estructural y dinámica muy detallada que nos permita esclarecer las bases moleculares que rigen estos procesos.

5. IMPORTANCIA DE LAS INTERACCIONES PROTEÍNA-PROTEÍNA. PROTEÍNAS INTRÍNECAMENTE DESORDENADAS

Durante los procesos de interacción de las proteínas con otras moléculas, se producen cambios conformacionales que, en algunos casos, suponen la pérdida de la estructura tridimensional de proteínas plegadas o la estructuración total o parcial de otras que se encuentran desplegadas de forma innata. Esta gran flexibilidad, muestra ventajas funcionales en una gran variedad de situaciones biológicamente relevantes (Collins et al. 2008). Las proteínas que presentan estas últimas características conformacionales, que se manifiestan en un cierto grado de desorden estructural, son las proteínas intrínsecamente desordenadas (IDPs), de las siglas en inglés “Intrinsically Disordered Protein” (Garner et al. 1998; Romero et al. 1998; Wright & Dyson 1999; Tompa 2002; Dyson & Wright 2005; Sickmeier et al. 2007; Dunker et al. 2008; Tompa 2012).

El hecho de que la cantidad de IDPs alcance proporciones elevadas en eucariotas superiores o que este tipo de proteínas esté particularmente presente en las interacciones que intervienen en funciones reguladoras y de señalización, ha colocado el estudio de las IDPs en el foco de las investigaciones en biología estructural.

5.1. Paradigma estructura-función

Desde que en 1894 Fischer (Fischer 1894) escribiera: “para ser lo más gráfico posible, me gustaría decir que enzima y sustrato tienen que encajar entre sí como una cerradura y una llave, para que así puedan ejercer los efectos químicos la una sobre la otra”, las investigaciones y el conocimiento acerca de la estructura de proteínas se ha incrementado radicalmente. Durante el siglo XX, muchos fueron los investigadores y estudios que enunciaban y aceptaban el hecho de que una estructura tridimensional altamente específica era un prerequisite para la función de la proteína (Mirsky & Pauling 1936; Sela et al. 1957; Kendrew et al. 1960; Blake et al. 1965; Anfinsen 1973). A partir de los resultados de Karush acerca de lo que denominó adaptabilidad configuracional (Karush 1950) y el ajuste inducido (Koshland 1958) surgieron las primeras evidencias de que cambios conformacionales significativos son los responsables de la función de muchas proteínas. Fue en los años 80, cuando se descubrió que segmentos de proteínas, que no estaban estructurados, jugaban importantes papeles en la función de las mismas. El descubrimiento de los segmentos de proteínas intrínsecamente desordenados (Alber 1982; Spolar & Record 1994; Lewis et al. 1996) es un ejemplo más de que el modelo inicial basado en un comportamiento estático no describe adecuadamente el universo de las interacciones moleculares.

5.2. Proteínas intrínsecamente desordenadas

Las proteínas intrínsecamente desordenadas (IDPs) son proteínas caracterizadas por la ausencia de una estructura terciaria estable cuando la proteína se encuentra como una cadena peptídica aislada en condiciones fisiológicas (Dyson & Wright 2005; Dunker et al. 2008).

Las proteínas y regiones intrínsecamente desordenadas (IDPs e IDRs) se parecen a los estados desnaturalizados, no plegados, de las proteínas globulares. A diferencia de éstos, las IDPs/IDRs, a menudo son responsables de funciones importantes relacionadas con la señalización o regulación, adaptabilidad, modificaciones post-traduccionales, interacción rápida con otros agentes, etc.

La base de datos DisProt (<http://www.disprot.org/>) contiene la mayor colección de IDPs disponibles en la actualidad. En la última revisión, esta base de datos contaba con 1513 regiones desordenadas correspondientes a 684 IDPs. Resulta evidente que DisProt contiene sólo una pequeña fracción de todas las IDPs/IDRs, ya que las predicciones bioinformáticas sugieren que el desorden estructural es abundante en varias especies, con una distribución desigual relacionada con el grado de evolución, mostrando una abundancia significativamente más alta en eucariotas que en procariotas.

6. MÉTODOS EXPERIMENTALES PARA EL ESTUDIO DE LAS INTERACCIONES

En la actualidad se dispone de muchos métodos para estudiar interacciones moleculares, cada uno de los cuales presenta ventajas e inconvenientes, sobre todo en cuanto a la sensibilidad y a la especificidad del método. Estos métodos se pueden dividir en dos grandes grupos: los métodos experimentales y los métodos teóricos. Dentro de los primeros tenemos los métodos bioquímicos y los biofísicos.

Las interacciones intermoleculares de biomoléculas se pueden describir a nivel de resolución atómica o globalmente, y se pueden estudiar desde el punto de vista cinético, mediante velocidades de reacción, y termodinámico, a través del equilibrio químico. Dentro de las técnicas de análisis termodinámico encontramos, entre otras, la espectroscopía de dicroísmo circular (CD), la fluorescencia intrínseca de algunos aminoácidos, la calorimetría isotérmica de titulación (ITC) y la calorimetría diferencial de barrido (DSC).

Otras técnicas experimentales que permiten obtener información estructural de las interacciones intermoleculares son: las espectroscopías de infra-rojo y Raman, la espectrometría de masas, la microscopía de fuerza atómica (AFM). Pero sin duda, las dos técnicas que más información proporcionan para el estudio y el análisis estructural de complejos a nivel de resolución atómica son la cristalografía de rayos-X y la Resonancia Magnética Nuclear, RMN.

7. LA RMN COMO MÉTODO PARA ESTUDIAR INTERACCIONES FUERTES Y TRANSITORIAS

Las interacciones entre biomoléculas pueden ser clasificadas por la fortaleza de su interacción, la cual a menudo es interpretada a partir de la constante de equilibrio de disociación, K_d , ($K_d = k_{off}/k_{on}$, donde k_{off} es la constante de velocidad de la reacción de disociación del complejo y k_{on} es la de la reacción de formación). Décadas de estudios han llevado a dividir las interacciones de proteínas en dos subclases: las interacciones fuertes con una K_d del orden de nM (Nooren & Thornton 2003) y las interacciones débiles con una K_d mayor, del orden de μ M. Mientras que se ha avanzado de forma apreciable en el conocimiento acerca de la naturaleza de los complejos de proteínas de alta afinidad, nuestra comprensión sobre las características estructurales de las interacciones débiles que involucran proteínas continúa siendo limitada. Esto es atribuible, por una parte a las dificultades técnicas que entraña su caracterización directa *in vitro* o *in vivo*. Por otro lado, la idea extendida hasta hace relativamente poco tiempo de que las interacciones débiles no pueden ocurrir en células vivas dadas las concentraciones tan bajas de proteína en las mismas ($<10^{-7}$ M), no ha favorecido su estudio. Sin embargo, las evidencias surgidas en los últimos años, que sugieren que este tipo de interacciones son fundamentales en diversos procesos biológicos, tales como: promover la activación/desactivación de la señal de transducción, los contactos reversibles entre células, el ensamblaje o desensamblaje de grandes complejos de señalización y el reconocimiento entre enzimas y sustrato; hacen vital el conocimiento de las características a nivel atómico de las interacciones débiles (Vaynberg & Qin 2006).

Métodos convencionales como la cristalografía de rayos-X, la resonancia plasmónica de superficie (SPR), la calorimetría isotérmica de titulación (ITC) y la RMN en estado sólido a menudo fallan en el estudio de interacciones débiles. Por el contrario, la RMN en solución ha resultado ser una herramienta particularmente útil para examinar este tipo de interacciones a resolución atómica y en condiciones cercanas a las fisiológicas (Wüthrich 1986; Clore & Gronenborn 1998; Qin et al. 2001; Zuiderweg 2002; Carrozzo et al. 2004).

A través de parámetros de RMN como los desplazamientos químicos, el acoplamiento escalar y las constantes dipolares residuales se pueden determinar cambios estructurales derivados de las interacciones. Los experimentos de relajación nos aportan datos de la dinámica y la flexibilidad conformacional del sistema libre y del complejo. Los métodos de difusión traslacional, permiten determinar el tamaño de los complejos que se forman. El

efecto Overhauser nuclear (NOE), que es la transferencia de polarización a través de la relajación cruzada intermolecular, permite obtener información de las distancias intermoleculares. Todos estos datos nos proporcionan información acerca de tres cuestiones importantes para la caracterización de las interacciones: (I) confirmar si el complejo se forma, (II) determinar el valor de la K_d y (III) determinar la estructura del complejo. De los parámetros mencionados anteriormente, el cartografiado por perturbación de los desplazamientos químicos, el efecto Overhauser nuclear y el acoplamiento dipolar residual, resultan de especial interés:

- Cartografiado por perturbación de los desplazamientos químicos (δ). Las frecuencias de resonancia, y por tanto los desplazamientos químicos de los átomos individuales tiene una fuerte dependencia del entorno químico local y debido a esto se pueden considerar como las huellas dactilares de las estructuras. Para el caso de proteínas, los patrones de desplazamientos químicos de ^{15}N y de los ^1H unidos directamente en los grupos amida, son especialmente sensibles a este respecto. Así, sus perturbaciones, resultado de la formación de complejos, proporcionan una herramienta muy sensible para realizar un cartografiado de la superficie de unión. Experimentalmente, esta superficie de unión puede ser identificada a partir de titulaciones en las que se monitorizan los cambios espectrales en los espectros ^1H - ^{15}N HSQC (heteronuclear single quantum correlation) si los complejos son pequeños (de peso molecular inferior a aproximadamente 15-20 kDa) o empleando las variantes TROSY (transverse relaxation-optimized spectroscopy) (Pervushin et al. 1997) si se trata de proteínas grandes. Para el caso de proteínas y péptidos sin marcaje isotópico, el cartografiado se realiza generalmente siguiendo las variaciones de los protones H_N y H_α .
- Efecto Overhauser nuclear (NOE). La principal aplicación del efecto NOE es para el cálculo de estructura de biomoléculas, ya que la intensidad del mismo es inversamente proporcional a la sexta potencia de la distancia que separa a los dos núcleos (protones) involucrados. El NOE también se ha utilizado ampliamente para el estudio de complejos, sobre todo de proteínas con proteínas, de ligandos de bajo peso molecular con proteínas y de complejos de ácidos nucleicos con proteínas. Los complejos que están fuertemente unidos, se comportarán respecto a este parámetro de la misma forma que las proteínas individuales (Williamson 2009). Cuando están unidos más débilmente aparecen efectos adicionales relacionados con la elevada movilidad y los regímenes de

difusión de spin de las moléculas pequeñas (Brand et al. 2005), así como con los equilibrios conformacionales.

- Acoplamiento dipolar residual (RDC). Este parámetro de RMN ha demostrado ser una valiosa herramienta ya que puede proporcionar restricciones de largo alcance para la determinación de la estructura molecular o para tener información cualitativa acerca de cambios en la orientación de enlaces concretos. Las restricciones que proporcionan se complementan con las restricciones de distancias obtenidas a través de los NOEs (Tjandra & Bax 1997). Esta complementariedad puede ser extendida a la geometría de los ligandos unidos a proteínas. Los valores de RDC de especial importancia para la determinación estructural de proteínas son los observados en sistemas de espín heteronucleares como los pares ^1H - ^{15}N o ^1H - ^{13}C , conectados a través de un enlace simple, e incluso en sistemas homonucleares ^1H - ^1H , como pares H_N - H_α .

8. PROBLEMÁTICA Y SOLUCIONES. MIMÉTICOS TANTO DE MEMBRANAS COMO DE PROTEÍNAS

El hecho de que una membrana biológica nativa pueda estar compuesta por cientos de lípidos diferentes (Dowhan 1997), y que las proteínas de membrana se purifican típicamente fuera de su membrana nativa y después deban ser reconstituidas, hace indispensable el empleo de sistemas modelos simplificados para estudiar los principios básicos de las interacciones proteína-lípido. Tales sistemas modelo permiten el análisis detallado de las propiedades de los lípidos y de cómo éstos influyen en la estructura y la dinámica de las porciones transmembrana. Permite, a su vez, saber cómo las proteínas de membrana están ancladas en la interfase lípido-agua y cómo la longitud y la composición de los segmentos transmembrana influyen en la organización y dinámica de los lípidos. Además, la caracterización de estos sistemas modelo ha demostrado ser útil para refinar el enfoque computacional que se estaba empleando hace unas décadas y desarrollar nuevas técnicas para el estudio de interacciones proteína-lípido (Killian & Nyholm 2006).

En muchos casos, las proteínas son reconstituidas en vesículas (o liposomas) que están compuestas de bicapas formadas sólo por uno o dos tipos de lípidos. El régimen dinámico, en los sistemas en los que se forman vesículas grandes, hace que prácticamente sólo puedan estudiarse por RMN de estado sólido. Por el contrario, los casos en los que se forman vesículas pequeñas, que se mueven a una velocidad suficiente, sí satisfacen *a priori* las condiciones para poder caracterizarlas por RMN en solución. Sin embargo, incluso las vesículas más pequeñas pueden formar agregados con pesos moleculares del orden de 10^6 Da o mayores, demasiado grandes para ser estudiados por RMN en solución (Henry & Sykes 1994). Por este motivo, las vesículas no son un buen medio para solubilizar proteínas de membrana purificadas para los estudios de RMN en solución. Para solventar este problema se ha recurrido al empleo de disolventes orgánicos y medios micelares o detergentes, que mimetizan también el entorno de membrana.

Mientras que las proteínas de membrana “desnudas” normalmente no son solubles en agua, en ocasiones se solubilizan en presencia de disolventes orgánicos o en mezclas de los mismos (tipo cloroformo, metanol, trifluoroetanol, hexafluoroisopropanol, isopropanol, etc.). Estos disolventes orgánicos que mimetizan entornos apolares, promueven generalmente la formación de la estructura secundaria en los péptidos que es análoga a la que adoptan en presencia de micelas. Como ejemplo de esta aproximación citaremos el uso del trifluoroetanol (TFE), que se ha venido empleando profusamente y con éxito en la

determinación estructural de péptidos antimicrobianos que interaccionan y desestabilizan membranas (Clore et al. 1986; Yi et al. 1996). También, en algunos casos, se emplean mezclas de solventes orgánicos, las cuales pueden interaccionar con los dominios transmembrana de las proteínas de membrana, solubilizándolas sin la necesidad de emplear un detergente. Hay que señalar que los disolventes orgánicos no producen, en general, el ensanchamiento de línea que puede ocurrir en presencia de micelas de detergentes, lo que en algunos casos los hace ser objeto de elección. Sin embargo, a pesar de que los disolventes orgánicos aplicados al estudio de estos sistemas han permitido un gran avance, está en discusión el hecho de que la estructura que se estabiliza en estos medios sea en todos los casos un reflejo de la estructura nativa (Sonnichsen et al. 1992). Un ejemplo de congruencia estructural es la que se refiere a los estudios de RMN del polipéptido hormona glucagón en los que se ha demostrado que la conformación adoptada es similar en micelas de DPC (Braun et al. 1983), en disolución acuosa y en mezclas de TFE (Clore et al. 1986).

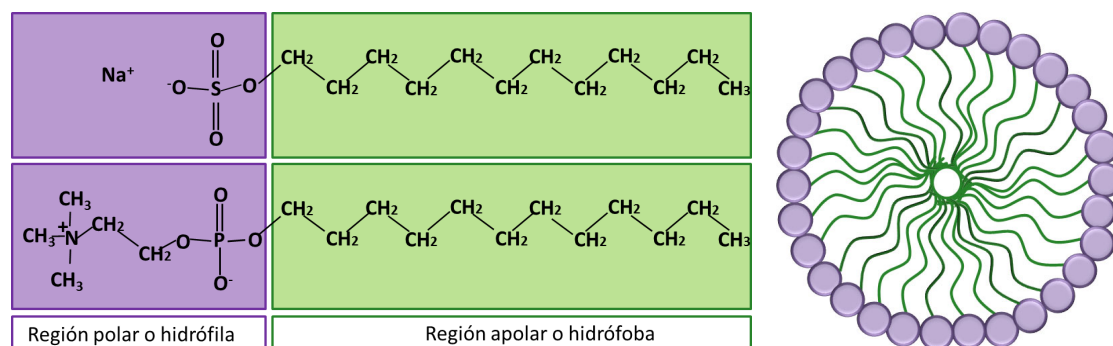


Figura 1.6. A la izquierda estructura de dos detergentes utilizados habitualmente en estudios de RMN en solución: en la parte superior el dodecilsulfato sódico (SDS) y en la parte inferior la dodecilmfosfocolina (DPC). A la derecha representación de una micela. La cabeza o región polar, que queda en contacto con solvente, está representada en morado y la cola o región apolar, que se sitúa hacia el interior de la micela, en verde.

Por otra parte, están los detergentes usados para solubilizar las proteínas de membrana, que generalmente se dividen en dos clases. La primera, la forman aquellos detergentes con colas apolares, figura 1.6. Dentro de este tipo encontramos el dodecilsulfato sódico (SDS), la dodecilmfosfocolina (DPC), N-lauoilsarcosina, óxido de laurildimetilamina (LDAO), etc. Los detergentes neutros presentan problemas ya que pueden conducir a espectros de RMN de baja calidad. Esto es debido a su alta monodispersabilidad que hace que las micelas no se mantengan bien a la elevada concentración que se requiere en los estudios de RMN. El uso de detergentes con cargas en los grupos de la cabeza está mucho más extendido y tienden a dar espectros de RMN de mayor calidad. Aun así, no hay una regla fija y cada sistema se comporta de manera diferente en presencia de los distintos medios (Krueger-Koplin et al. 2004). Además, detergentes como el SDS, N-lauoilsarcosina, DPC y LDAO tienen que ser

usados con cuidado por su propensión a desestabilizar proteínas (Sanders & Sonnichsen 2006). Mucho mejores son los fosfolípidos de cadena corta tipo dihexanoilfosfatidilcolina (DHPC) (Hauser 2000) y los lisofosfolípidos (Krueger-Koplin et al. 2004). La carga zwitteriónica en estos detergentes confiere alta solubilidad en agua y monodispersibilidad a sus micelas. A la vez, sus efectos sobre las proteínas de membrana son muy suaves. La transición de polar a apolar dentro de estos detergentes es gradual, justo como lo es la transición de lípidos en las bicapas.

Una segunda clase de detergentes biológicos la componen moléculas anfipáticas rígidas como péptidos helicoidales anfipáticos y detergentes derivados de sales biliares como por ejemplo 3-[(3-colamidopropil)dimetilamonio]-1-propanosulfonato sódico (CHAPS). Además de su utilidad para mejorar la solubilización de las proteínas de membrana, CHAPS se caracteriza por su capacidad para estabilizar partes hidrófobas expuestas de algunas proteínas solubles en agua, de este modo aumenta su solubilidad, algo que resulta muy útil para muestras que han de estar muy concentradas para estudios estructurales (McGuire et al. 1998). Gracias a la experiencia y a los estudios publicados en la literatura, sabemos que los detergentes de sales biliares son muy buenos para solubilizar complejos lípidos/proteínas.

9. OBJETIVO GENERAL

A lo largo de esta introducción se ha hablado de la importancia de las interacciones entre biomacromoléculas, en especial las que se dan entre proteínas y lípidos, y del papel fundamental que desempeña la RMN como técnica en la caracterización de las mismas. A partir de estas dos premisas, el trabajo que se presenta en esta Tesis tiene dos objetivos bien diferenciados.

El primero pretende, desde un punto de vista metodológico, explorar las limitaciones de la aplicación de la RMN a las proteínas carentes de estructura. Como se ha expuesto, estas proteínas están presentes en muchas interacciones que intervienen en funciones vitales. Por otra parte, aunque el mecanismo por el cual una proteína atraviesa una membrana no es bien conocido desde el punto de vista estructural, se ha postulado que puede ocurrir a través de estados desplegados o parcialmente desplegados. Por ello, en el contexto de esta Tesis, se ha creído interesante aportar una nueva herramienta que facilita la correcta e inambigua asignación espectral de estos sistemas altamente dinámicos. Ésta puede ser especialmente útil en casos en los que las aproximaciones actuales no resultan totalmente satisfactorias.

Por otro lado, y desde un punto de vista aplicado, se ha pretendido aportar nueva información acerca de los mecanismos, factores determinantes de la estructura y funcionamiento de ciertas proteínas que interaccionan con membranas. Para ello, empleando la RMN en disolución, se han estudiado las interacciones que se establecen así como los procesos de fusión de las membranas en varios sistemas, cuyo entendimiento nos podría ayudar a diseñar agentes terapéuticos eficaces para el tratamiento de distintas enfermedades.

Esta memoria se ha estructurado en cinco capítulos que engloban los distintos trabajos realizados durante este periodo. En el presente capítulo se ha realizado una introducción en la que se detalla la importancia de las interacciones entre biomacromoléculas y la relevancia de la RMN para la caracterización estructural y el estudio a resolución atómica de las mismas. En el capítulo 2, se aborda la RMN desde un punto de vista metodológico y se describe una nueva secuencia de detección de $^{13}\text{C}_\alpha$. En el capítulo 3, se tratan de determinar las interacciones que se producen entre los péptidos β -boomerang y un derivado de la heparina empleado en medicina clínica, el Fondaparinux. En el capítulo 4, se estudian distintas regiones del ectodominio de la glicoproteína gp41 de VIH-1 en medios que mimetizan el entorno de membranas, como son el HFIP y el DPC. En el capítulo 5, se realiza la determinación estructural del zimógeno ONCFLG de la Onconasa® que nos permitirá

comprender las bases estructurales y dinámicas de la actividad catalítica y la alta estabilidad de esta enzima en sus formas completa e hidrolizada como paso previa al estudio de su interacción con membranas. Por último, en el capítulo 6, se exponen brevemente las conclusiones generales de esta Tesis doctoral.

10. BIBLIOGRAFÍA

- Alber T, Gilbert, W.A., Ponzi, D.R., and Petsko, G.A. (1982) The role of mobility in the substrate binding and catalytic machinery of enzymes. *Ciba Found Symposium* 93, 4-24.
- Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181, 223-230.
- Blake CC, Koenig DF, Mair GA et al. (1965) Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* 206, 757-761.
- Brand T, Cabrita EJ, Berger S (2005) Intermolecular interaction as investigated by NOE and diffusion studies. *Prog NMR Spectroscopy* 46, 159-196.
- Braun W, Wider G, Lee KH et al. (1983) Conformation of glucagon in a lipid-water interphase by ¹H nuclear magnetic resonance. *Journal of molecular biology* 169, 921-948.
- Buckle AM, Schreiber G, Fersht AR (1994) Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-Å resolution. *Biochemistry* 33, 8878-8889.
- Cai YH, Huang H (2012) Advances in the study of protein-DNA interaction. *Amino acids* 43, 1141-1146.
- Carrozzo R, Rizza T, Lucoli S et al. (2004) A mitochondrial ATPase 6 mutation is associated with Leigh syndrome in a family and affects proton flow and adenosine triphosphate output when modeled in *Escherichia coli*. *Acta Paediatr Suppl* 93, 65-67.
- Case DA, Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S. and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.
- Cavanagh J, Fairbrother WJ, Palmer III AG et al. (1996) Protein NMR spectroscopy: Principles & Practice. Academic Press.
- Clore GM, Gronenborn AM (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends in biotechnology* 16, 22-34.
- Clore GM, Martin SR, Gronenborn AM (1986) Solution structure of human growth hormone releasing factor. Combined use of circular dichroism and nuclear magnetic resonance spectroscopy. *Journal of molecular biology* 191, 553-561.
- Collins MO, Yu L, Campuzano I et al. (2008) Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Molecular & cellular proteomics : MCP* 7, 1331-1348.
- Cosic I (1994) Macromolecular bioactivity: is it resonant interaction between macromolecules?--Theory and applications. *IEEE transactions on bio-medical engineering* 41, 1101-1114.
- Chen JC, Miercke LJ, Krucinski J et al. (1998) Structure of bovine pancreatic cholesterol esterase at 1.6 Å: novel structural features involved in lipase activation. *Biochemistry* 37, 5107-5117.
- Dowhan W (1997) Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annual review of biochemistry* 66, 199-232.
- Dunker AK, Silman I, Uversky VN et al. (2008) Function and structure of inherently disordered proteins. *Current opinion in structural biology* 18, 756-764.
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nature reviews Molecular cell biology* 6, 197-208.
- Engelman DM (2005) Membranes are more mosaic than fluid. *Nature* 438, 578-580.
- Fischer E (1894) Einfluss der configuration auf die wirkung der enzyme. *Ber Dt Chem Ges* 27, 2985-2993.

- Garner E, Cannon P, Romero P et al. (1998) Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. *Genome informatics Workshop on Genome Informatics* 9, 201-213.
- Gonen T, Cheng Y, Sliz P et al. (2005) Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438, 633-638.
- Graham DL, Lowe PN, Grime GW et al. (2002) MgF(3)(-) as a transition state analog of phosphoryl transfer. *Chemistry & biology* 9, 375-381.
- Han X, Gross RW (2005) Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass spectrometry reviews* 24, 367-412.
- Hauser H (2000) Short-chain phospholipids as detergents. *Biochimica et biophysica acta* 1508, 164-181.
- Henry GD, Sykes BD (1994) Methods to study membrane protein structure in solution. *Methods in enzymology* 239, 515-535.
- Hunte C (2005) Specific protein-lipid interactions in membrane proteins. *Biochemical Society transactions* 33, 938-942.
- Kami K, Rajesh S, Overduin M (2009) Phospholipid-interacting proteins by solution-state NMR spectroscopy. *Methods Mol Biol* 462, 291-306.
- Karush F (1950) Heterogeneity of the Binding Sites of Bovine Serum Albumin. *Journal of the American Chemical Society*, Vol. 72. pp. 2705-2713.
- Kendrew JC, Dickerson RE, Strandberg BE et al. (1960) Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å. resolution. *Nature* 185, 422-427.
- Killian JA, Nyholm TK (2006) Peptides in lipid bilayers: the power of simple models. *Current opinion in structural biology* 16, 473-479.
- Kolatkhar AR, Weis WI (1996) Structural basis of galactose recognition by C-type animal lectins. *The Journal of biological chemistry* 271, 6679-6685.
- Koshland JDE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 44, 98-104.
- Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST et al. (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *Journal of biomolecular NMR* 28, 43-57.
- Lewis M, Chang G, Horton NC et al. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271, 1247-1254.
- Marsh D, Pali T (2004) The protein-lipid interface: perspectives from magnetic resonance and crystal structures. *Biochimica et biophysica acta* 1666, 118-141.
- McGuire AM, Matsuo H, Wagner G (1998) Internal and overall motions of the translation factor eIF4E: cap binding and insertion in a CHAPS detergent micelle. *Journal of biomolecular NMR* 12, 73-88.
- Mirsky AE, Pauling L (1936) On the Structure of Native, Denatured, and Coagulated Proteins. *Proceedings of the National Academy of Sciences of the United States of America* 22, 439-447.
- Monod J, Wyman J, Changeux JP (1965) On the Nature of Allosteric Transitions: A Plausible Model. *Journal of molecular biology* 12, 88-118.
- Nooren IM, Thornton JM (2003) Diversity of protein-protein interactions. *The EMBO journal* 22, 3486-3492.
- Palsdottir H, Hunte C (2004) Lipids in membrane protein structures. *Biochimica et biophysica acta* 1666, 2-18.
- Park SH, Mrse AA, Nevzorov AA et al. (2006) Rotational diffusion of membrane proteins in aligned phospholipid bilayers by solid-state NMR spectroscopy. *J Magn Reson* 178, 162-165.
- Perkins JR, Diboun I, Dessailly BH et al. (2010) Transient protein-protein interactions: structural, functional, and network properties. *Structure* 18, 1233-1243.

- Pervushin K, Riek R, Wider G et al. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12366-12371.
- Qin J, Vinogradova O, Gronenborn AM (2001) Protein-protein interactions probed by nuclear magnetic resonance spectroscopy. *Methods in enzymology* 339, 377-389.
- Romero P, Obradovic Z, Kissinger CR et al. (1998) Thousands of proteins likely to have long disordered regions. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*, 437-448.
- Ruff M, Krishnaswamy S, Boeglin M et al. (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA(Asp). *Science* 252, 1682-1689.
- Sanders CR, Sonnichsen F (2006) Solution NMR of membrane proteins: practice and challenges. *Magnetic resonance in chemistry : MRC 44 Spec No*, S24-40.
- Sela M, White FH, Jr., Anfinsen CB (1957) Reductive cleavage of disulfide bridges in ribonuclease. *Science* 125, 691-692.
- Sickmeier M, Hamilton JA, LeGall T et al. (2007) DisProt: the Database of Disordered Proteins. *Nucleic acids research* 35, D786-793.
- Singer SJ, Nicolson GL (1972) The fluid mosaic model of the structure of cell membranes. *Science* 175, 720-731.
- Sonnichsen FD, Van Eyk JE, Hodges RS et al. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* 31, 8790-8798.
- Spolar RS, Record MT, Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science* 263, 777-784.
- Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* 278, 1111-1114.
- Tompa P (2002) Intrinsically unstructured proteins. *Trends in biochemical sciences* 27, 527-533.
- Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. *Trends in biochemical sciences* 37, 509-516.
- Tycko R (2001) Biomolecular solid state NMR: advances in structural methodology and applications to peptide and protein fibrils. *Annual review of physical chemistry* 52, 575-606.
- Vaynberg J, Qin J (2006) Weak protein-protein interactions as probed by NMR spectroscopy. *Trends in biotechnology* 24, 22-27.
- von Heijne G, Rees D (2008) Membranes: reading between the lines. *Current opinion in structural biology* 18, 403-405.
- Williamson MP (2009) Applications of the NOE in Molecular Biology. *Annual Reports on NMR Spectroscopy* 65.
- Wright PE, Dyson HF (1999) Intrinsically unstructured proteins: re assesseng the protein structure-function paradigm. *Journal of molecular biology* 293, 321-331.
- Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. J Wiley&Sons, Inc New York.
- Yi GS, Park CB, Kim SC et al. (1996) Solution structure of an antimicrobial peptide buforin II. *FEBS letters* 398, 87-90.
- Zuiderweg ER (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* 41, 1-7.

Capítulo II

Desarrollo de la secuencia
inter-CAN, de detección de ^{13}C , para el
estudio de estados desplegados

Capítulo II:

Desarrollo de la secuencia inter-CAN, de detección de ^{13}C , para el estudio de estados desplegados

1. INTRODUCCIÓN

Como se ha mencionado en el capítulo I, la RMN en disolución juega un papel estratégico en la caracterización y determinación estructural de macromoléculas y sus complejos, ya que permite el acceso a la información estructural y dinámica a resolución atómica de las mismas.

En este aspecto, el gran éxito de la RMN se debe a su versatilidad y posibilidad para realizar una amplia gama de experimentos, permitiendo así el desarrollo de diferentes secuencias de pulsos que se pueden aplicar a una gran variedad de sistemas.

Hoy en día, la espectroscopía de RMN es todavía una técnica en rápido crecimiento, que continúa aumentando sus aplicaciones, especialmente en química, bioquímica y medicina. El avance de la espectroscopía de RMN sigue dos líneas principales: una centrada en la tecnología del equipo y la otra basada en el desarrollo de nuevas técnicas experimentales. En cuanto a lo que concierne a la primera, los esfuerzos se han centrado en el desarrollo de campos magnéticos más altos y en dotar a los espectrómetros con sondas de alta calidad, pre-amplificadores, generadores de campos de radiofrecuencia y ordenadores modernos que nos permiten realizar las medidas con mayor sensibilidad y facilitan la aplicación de los

nuevos métodos experimentales. En cuanto al progreso metodológico, las técnicas experimentales mejoran de forma continua con el fin de aumentar la sensibilidad, la precisión en la determinación experimental o la detección de nuevos observables. Las nuevas secuencias de pulsos de RMN que se diseñan constantemente, forman parte de este último grupo de avances y es el marco en el que se desarrolla este capítulo.

1.1. Problemas de la detección de ^1H y alternativas

El hecho de que el núcleo ^1H posea el valor más alto de constante giromagnética ($\gamma^1\text{H} \approx 26.7 \times 10^7$) entre todos los elementos de la tabla periódica (a excepción del tritio que es muy inestable y radiactivo), le proporciona una alta sensibilidad, por lo que es el núcleo más usado para ser observado por RMN. Sin embargo, la RMN de detección de ^1H está limitada principalmente por la elevada contribución dipolar a la relajación. Ello tiene como consecuencia la rápida relajación transversal provocando el ensanchamiento de señales y la reducción de la resolución espectral. Esto es un problema especialmente importante en la adquisición de espectros multidimensionales de biomoléculas.

A lo largo de los últimos años se han desarrollado distintas herramientas para solventar estos problemas. Entre ellas cabe destacar el uso del fenómeno de relajación cruzada “cross-correlated”, que reduce la velocidad de la relajación transversal efectiva de espines específicos (Pervushin et al. 1997; Pervushin et al. 1998; Tugarinov et al. 2003; Miclet et al. 2004) y el enriquecimiento isotópico con ^2H , cuyo menor valor de constante giromagnética y su efecto de “dilución” de la población de ^1H reduce drásticamente las interacciones dipolo-dipolo, produciendo mecanismos de relajación menos eficientes (LeMaster & Richards 1988; Markus et al. 1994; Sattler & Fesik 1996). Sin embargo, en algunos casos estas estrategias no son suficientes y hay que buscar alternativas que solventen los problemas planteados. Una de estas soluciones es la detección directa de otros núcleos. Hoy en día, para el estudio de biomoléculas sólo se emplea la detección de ^{13}C ($\gamma^{13}\text{C} \approx 6.72 \times 10^7$), ya que en el caso del ^{15}N la constante giromagnética es demasiado pequeña ($\gamma^{15}\text{N} \approx 2.71 \times 10^7$) y por ende también la sensibilidad que puede alcanzarse, figura 2.1.

1.2. Detección directa de ^{13}C

La detección directa de ^{13}C implica la adquisición de las frecuencias de los átomos de ^{13}C en la dimensión de medida de un experimento de RMN. Actualmente, gracias a la mejora instrumental, al uso de campos magnéticos más altos y de criosondas más sensibles, la detección directa de ^{13}C se ha convertido en un recurso valioso para los espectroscopistas de RMN. La espectroscopía de detección directa de ^{13}C puede ser empleada tanto en solución (Bermel et al. 2006c), como en estado sólido (Fyfe 1983) e incluso en casos muy particulares como la RMN en células vivas (Selenko & Wagner 2007).

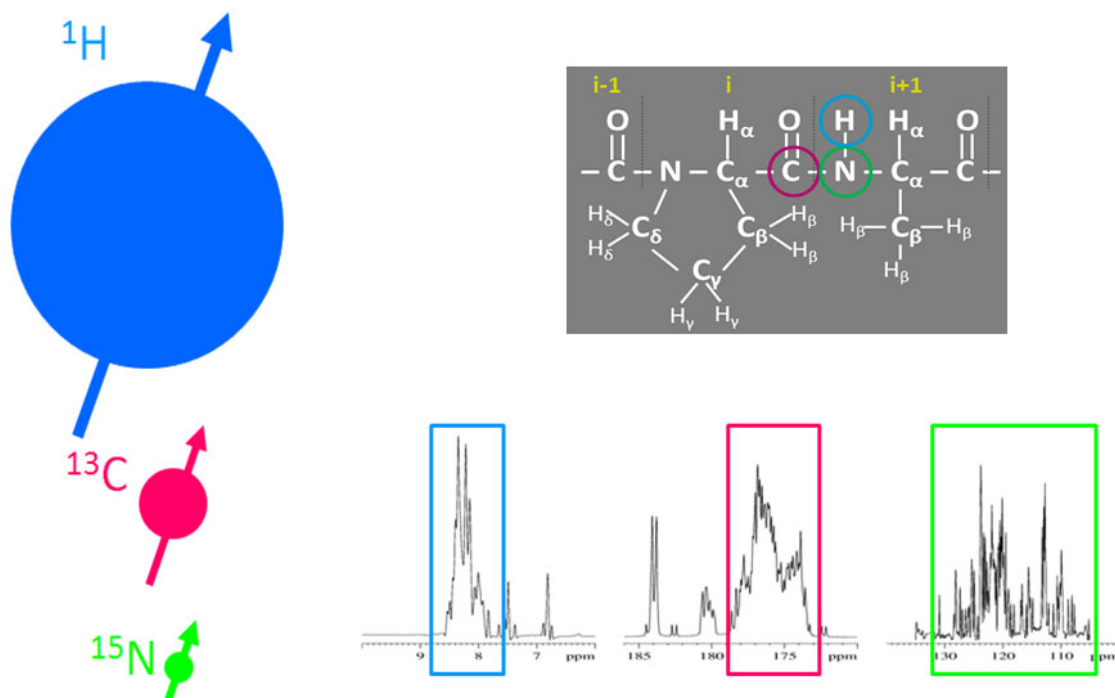


Figura 2.1. Diferencia en las propiedades intrínsecas de los núcleos ^1H , ^{13}C y ^{15}N . A la izquierda, una representación, a escala, de los diferentes valores de constante giromagnética, γ , responsable de la baja sensibilidad del ^{13}C y ^{15}N con respecto al ^1H y de la reducción de las contribuciones a la relajación dipolar ($\gamma^1\text{H} \approx 26.7 \times 10^7$, $\gamma^{13}\text{C} \approx 6.72 \times 10^7$, $\gamma^{15}\text{N} \approx -2.71 \times 10^7$). A la derecha, incremento de la dispersión de los desplazamientos químicos de ^1H , ^{13}C y ^{15}N en algunas zonas de interés para proteínas. En azul los ^1H , en rosa los ^{13}C y verde los ^{15}N . Esta imagen ha sido cedida por la Profesora Isabella Caterina Felli del Departamento de Química de la Universidad de Florencia, CERM, Italia.

Aunque la detección directa de ^{13}C se ha empleado desde los comienzos de la RMN en el estudio de moléculas pequeñas, como compuestos orgánicos e inorgánicos (Harris 1986), o metabolitos (Lindon et al. 2005), no ha sido utilizada de forma habitual en el estudio de macromoléculas biológicas debido a la baja sensibilidad que presenta en comparación con el protón. Sin embargo, en la actualidad el uso extensivo de moléculas marcadas isotópicamente en ^{13}C ha favorecido la expansión de esta metodología (Grzesiek & A. 1992).

En principio cualquier núcleo de ^{13}C en las biomoléculas podría usarse para ser observado. Sin embargo, motivos técnicos y de eficiencia para la asignación de espectros de proteínas, han propiciado el desarrollo de secuencias de pulsos dedicadas a la detección de los átomos del esqueleto, C' (se designan así los carbonos correspondientes al grupo carbonilo del enlace peptídico) y de la posición alfa, C $_{\alpha}$.

1.2.1. Características de la detección directa de ^{13}C

Las ventajas del uso de la RMN de detección directa de ^{13}C para biomoléculas marcadas isotópicamente se derivan de las propiedades intrínsecas de este núcleo. Concretamente, la reducida contribución dipolar a la relajación, el incremento de la dispersión de los desplazamientos químicos, y el hecho de que los núcleos no se ven afectados por procesos de intercambio químico con el solvente, hacen de la RMN de detección directa de ^{13}C la elección natural para el desarrollo de experimentos que nos permitan el estudio de proteínas allí donde la detección de ^1H se vería abocada al fracaso.

Además, la detección de ^{13}C presenta otras ventajas. Entre ellas, el hecho de que puede obtenerse información de los carbonos cuaternarios, por ejemplo carbonilos, y de todos los carbonos en muestras deuteradas sin necesidad de recurrir a múltiples pasos de transferencia como ocurre en el caso de la detección de los protones H $_N$. Por último, la detección de ^{13}C evita el problema de la eliminación de la señal del agua.

A pesar de las ventajas, la detección de ^{13}C presenta, también, algunos inconvenientes como son: la disminución en la sensibilidad y la necesidad de eliminar el acoplamiento homonuclear ^{13}C - ^{13}C , que se exponen a continuación.

1.2.1.1. Sensibilidad

En un experimento de RMN, la sensibilidad o relación señal/ruido puede expresarse por la siguiente ecuación:

$$S/N \propto \gamma_{ex} \cdot \gamma_{det}^{3/2} \cdot B_0^{3/2} \cdot N^{1/2}$$

Donde γ_{ex} y γ_{det} son respectivamente las constantes giromagnéticas de los núcleos excitados al comienzo de la secuencia y los detectados al final de la misma. B_0 , es la intensidad del campo magnético del espectrómetro utilizado y N es el número de barridos del experimento de RMN.

De la ecuación anterior se deduce que, sin considerar los efectos de relajación, cuanto mayor sea la constante giromagnética, γ , del núcleo excitado y detectado, mayor será la

sensibilidad. Por este motivo, la detección directa de ^{13}C ha sido tradicionalmente problemática dado que su constante giromagnética es prácticamente un cuarto de la del ^1H ($\gamma_{\text{C}}/\gamma_{\text{H}} = 1/4$), dando lugar a una sensibilidad de $1/64$ con respecto al ^1H .

Por tanto, con el fin de conseguir la máxima sensibilidad en los experimentos de detección directa de ^{13}C , es importante optimizar los núcleos sobre los que se actúa y en los que comienza la polarización, independientemente del protocolo usado para la transferencia de coherencia y su eficiencia relativa. Ello se traduce en que, generalmente, primero se excitan los protones (lo que permite aprovechar las ventajas de su gran valor de γ), y después se detecta el ^{13}C .

1.2.1.2. Resolución y anchura de línea

El valor más bajo de constante giromagnética de ^{13}C , con respecto al ^1H , reduce drásticamente las interacciones dipolo-dipolo produciendo mecanismos de relajación menos eficientes. En este contexto, la velocidad de relajación transversal, que condiciona las anchuras de línea en los experimentos de RMN, es uno de los factores ventajosos en la detección de ^{13}C . A altos campos magnéticos y en sistemas deuterados, los carbonos alifáticos son los núcleos que tienen propiedades de relajación óptimas para el estudio de biomoléculas (Robert 1995). En ellos, la contribución más importante a la relajación transversal es la relajación dipolo-dipolo, que sólo depende marginalmente de la intensidad del campo magnético. Esta ventaja se puede explotar para tener mejor sensibilidad y registrar espectros con señales más finas. Sin embargo, la detección de carbonos alifáticos, como los C_α , presenta el inconveniente de que requiere el desacoplamiento homonuclear ^{13}C - ^{13}C , figura 2.1.

En el caso de la detección de $^{13}\text{C}'$, el efecto dominante de la relajación transversal lo constituye la anisotropía de desplazamiento químico, que depende en gran medida de la intensidad del campo magnético B_0 (Pervushin et al. 1997). A pesar de que las secuencias de pulsos basadas en este núcleo se han desarrollado desde hace ya un tiempo debido a ventajas técnicas (Bertini et al. 2004a; Kovacs et al. 2005a), la detección del C' no es óptima para sistemas grandes y altos campos (Sattler et al. 1999) ya que puede conducir a espectros con señales anchas, resultado de la gran anisotropía de los desplazamientos químicos de C' , que provoca mecanismos de relajación muy eficientes que aumentan con el cuadrado del campo magnético.

Por otra parte, el hecho de que el intervalo de desplazamientos químicos para el ^{13}C de la mayor parte de compuestos sea de unas 200 ppm, mucho mayor al del protón que es de unas 15 ppm, tiene como consecuencia el menor solapamiento de las señales en los espectros de ^{13}C que en los de ^1H . Por ello puede conseguirse una mayor resolución espectral cuando empleamos la detección de ^{13}C .

1.2.1.3. Acoplamiento homonuclear

El fuerte acoplamiento homonuclear ^{13}C - ^{13}C es un factor beneficioso en determinados aspectos puesto que facilita la transferencia de coherencias en los experimentos de RMN. De hecho se hace un uso extensivo de esta propiedad tanto en experimentos de detección de ^1H como en ^{13}C . Sin embargo, también supone el principal problema al que debemos enfrentarnos en la detección directa de ^{13}C dado que causa el desdoblamiento indeseado de las señales. Por ejemplo, los C' se acoplan con los espines de la posición α (C_α) produciendo un aumento en el número de señales en los espectros y afectando a la resolución. Más complejo aún es el caso de los C_α que se acoplan a su vez a los C' y a los C_β (excepto las Gly que carecen de este último núcleo), figura 2.1, lo que conduce a multipletes complejos. Para evitar este efecto, y conseguir el colapso del multiplete en una única señal, se han propuesto varias soluciones en la literatura (Bermel et al. 2005b; Merlino et al. 2005; Bermel et al. 2006c; Bermel et al. 2007).

Considerando las distintas alternativas, en este trabajo se ha empleado la aproximación tipo IPAP ("In-Phase Anti-Phase"), figura 2.2 (A) y (B), (Duma et al. 2003; Bermel et al. 2007) que puede ser aplicada al estudio de una gran mayoría de los sistemas analizables por RMN, especialmente los sistemas desordenados dado que sus valores de relajación transversal son compatibles con esta implementación. En la aproximación IPAP, el reenfoque del desdoblamiento se realiza mediante el registro de dos FIDs para cada incremento, uno para la componente en-fase y otro para la componente en anti-fase. Después, ambos se suman y restan dando lugar a las dos componentes del multiplete que, finalmente, son desplazados en $J_{\text{C}'-\text{C}_\alpha}/2$ Hz hacia el centro del multiplete original y sumados para obtener un singlete como muestra la figura 2.2 (C). De esta manera, utilizando esquemas de pulsos selectivos, el desdoblamiento C' - C_α en la dimensión directa se puede eliminar de manera efectiva.

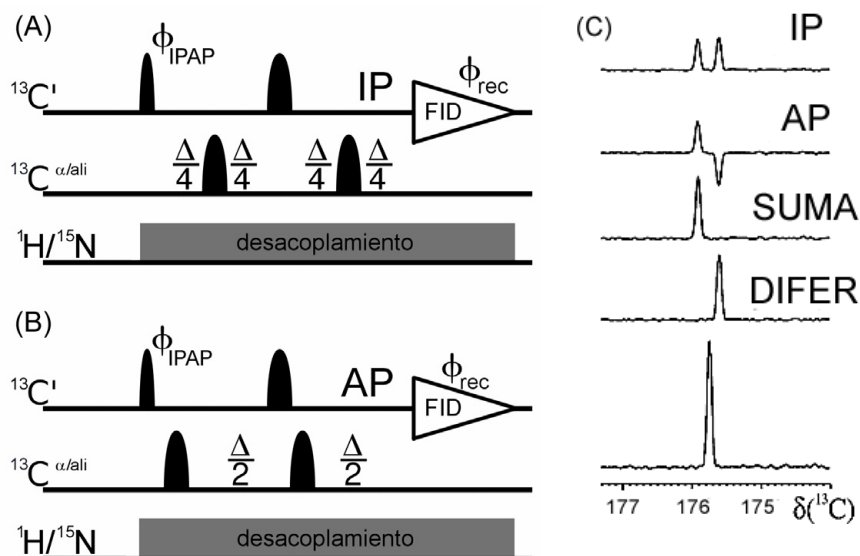


Figura 2.2. Aproximación IPAP para eliminar el desdoblamiento $\text{C}'\text{-C}_\alpha$ en detección directa de $^{13}\text{C}'$. (A) variante de la secuencia para el registro de la componente en-fase, IP. (B) Variante de la secuencia para el registro de la componente en anti-fase, AP. Aunque están representados para el caso de los experimentos 1D, pueden ser implementados en cualquier experimento basado en la detección directa de ^{13}C . (C) Representación esquemática del resultado de los dos experimentos (A) y (B) así como de su suma y diferencia y del espectro final resultado de la combinación de ambos. Los pulsos selectivos de ^{13}C de 90° son los estrechos y los de 180° los anchos. El tiempo de espera Δ es $1/(2J_{\text{C}_\alpha\text{C}'})$ (9ms).

En el caso de la detección directa de C_α (aproximación que se usa en la secuencia desarrollada en este trabajo), ya hemos comentado que la situación es más complicada. En principio, el desacoplamiento conjunto de los núcleos de C' y de C_β se puede lograr con las mismas estrategias empleadas para la detección de C' . En el caso de C_α , esto se consigue combinando dos bloques IPAP, uno específico para el acoplamiento $\text{C}_\alpha\text{-C}'$ (55 Hz) y otro para el $\text{C}_\alpha\text{-C}_\beta$ (35 Hz). De hecho, los dos bloques IPAP pueden ser concatenados en un bloque más corto cuya duración viene determinada por el valor de la constante de acoplamiento más pequeña de entre las involucradas. La figura 2.3 muestra la doble implementación IPAP (DIPAP de aquí en adelante) que ha sido propuesta para el desacoplamiento en la detección directa de C_α .

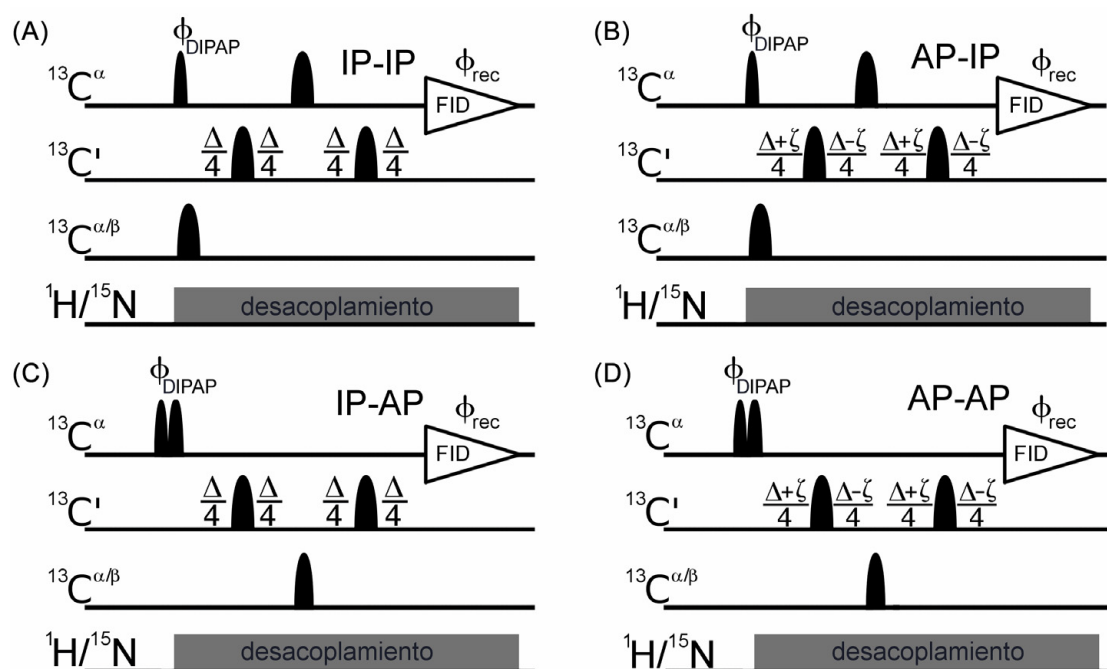


Figura 2.3. Método DIPAP para detección directa de C_α para eliminar los acoplamientos $\text{C}_\alpha\text{-C}'$ y $\text{C}_\alpha\text{-C}_\beta$ en la dimensión directa de adquisición. Las imágenes muestran el caso más simple de un experimento 1D.

1.2.2. Aplicaciones de la detección directa de ^{13}C para biomoléculas

Hasta hace poco tiempo los experimentos de detección directa de ^{13}C apenas se habían empleado en los estudios de macromoléculas biológicas debido a los problemas mencionados anteriormente. Gracias a las mejoras tecnológicas alcanzadas en la actualidad, la sensibilidad de la detección de ^{13}C ha aumentado en un orden de magnitud, permitiendo el uso de la misma en muestras marcadas isotópicamente (Bertini et al. 2004a; Kovacs et al. 2005b). Como se puede observar en el espectro 1D de ^{13}C de la ubiquitina, figura 2.4, las señales de ^{13}C para los distintos tipos de carbono están bien separadas y pueden proporcionar información detallada del esqueleto y las cadenas laterales.

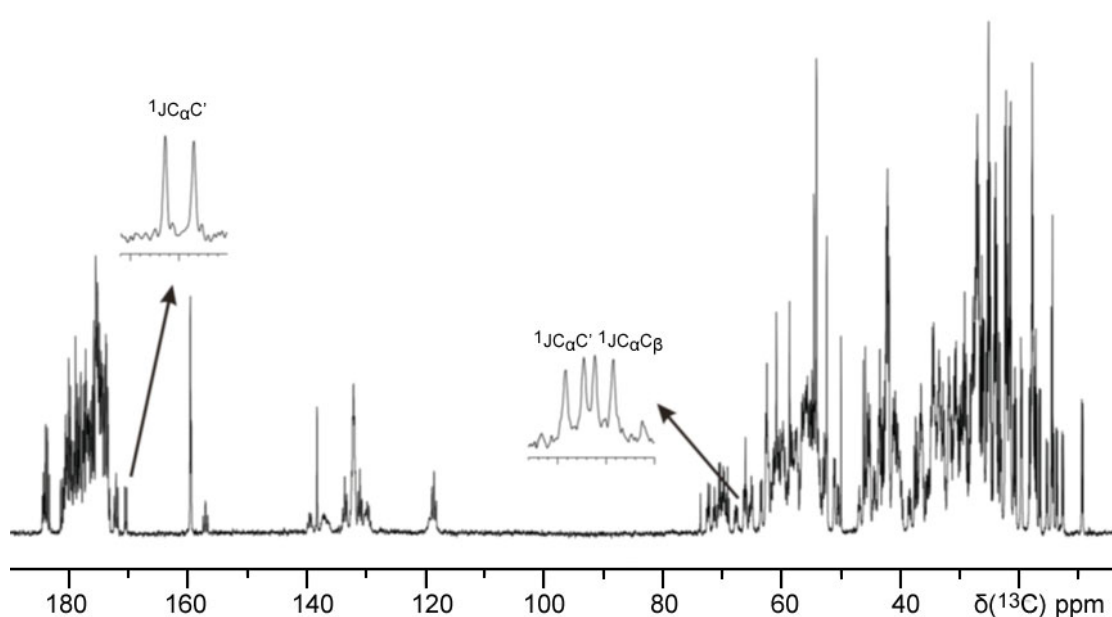


Figura 2.4. Espectro 1D de ^{13}C de una muestra 0.5 mM de ubiquitina (modelo clásico de proteína plegada empleada en los experimentos de RMN) doblemente etiquetada en ^{13}C y ^{15}N adquirida en un espectrómetro Bruker Avance equipado con una criosonda TXI. El espectro se adquirió con 64 barridos usando una secuencia de pulsos estándar con desacoplamiento de ^1H y ^{15}N durante la adquisición.

Recientemente, se han desarrollado numerosos experimentos, exclusivamente heteronucleares basados en la detección de $^{13}\text{C}'$, para la identificación de los sistemas de espín y la asignación secuencial, así como para la determinación de los observables de RMN más informativos (desplazamientos químicos, constantes de acoplamiento escalar, constantes dipolares residuales, NOEs, velocidades de relajación de ^{15}N) (Bermel et al. 2008). Estos experimentos son complementarios a aquellos basados en la detección de ^1H (Wüthrich 1986; Oschkinat et al. 1994), de manera que pueden proporcionar datos en situaciones donde la detección de protón es problemática. Los casos en los que la detección de ^{13}C resulta particularmente interesante y prácticamente indispensable son: proteínas paramagnéticas, sistemas de gran tamaño y eficiente relajación de espín, proteínas/regiones intrínsecamente o altamente desestructuradas y sistemas con intercambio rápido con el solvente.

1.2.2.1. Proteínas paramagnéticas

La contribución a la relajación nuclear depende del cuadrado de la constante giromagnética del núcleo observado. En el caso de proteínas que contengan un núcleo paramagnético los mecanismos eficaces de su relajación hacen que los núcleos, como el ^1H , con alto valor de γ sufran un gran ensanchamiento de señales, mientras que los núcleos con bajo valor de γ tales como ^{13}C y ^{15}N (Bertini et al. 2001; Bertini et al. 2005), se vean menos afectados. Por tanto, la detección de ^{13}C es aconsejable en estos casos ya que nos permite obtener información en regiones de proteínas paramagnéticas donde las anchuras de línea de ^1H estarían por encima de los límites detectables.

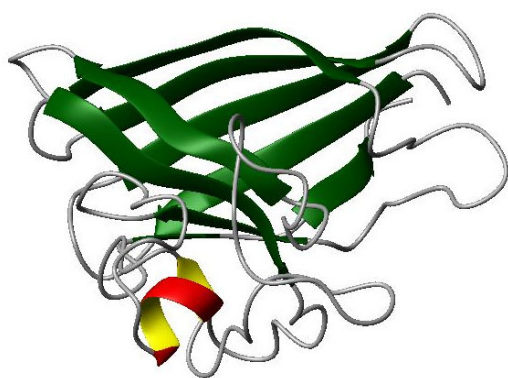


Figura 2.5. Proteína paramagnética. Monómero de cobre cinc superóxido dismutasa, formada por 153 aminoácidos. Estructura de RMN, código pdb 1DSW (Banci et al. 1999).

1.2.2.2. Sistemas de gran tamaño y eficiente relajación de espín

En proteínas grandes (con pesos moleculares superiores a 30 kDa (Lee et al. 2006)), cuyo tiempo de correlación rotacional en disolución es muy largo, se producen procesos de relajación transversal muy eficientes que, como ya se ha comentado, llegan a ser el factor limitante para los experimentos de ^1H -RMN en disolución. Los experimentos comúnmente empleados para la asignación específica y secuencial en estos casos fallan. El uso del fenómeno de relajación cruzada (Kumar et al. 2000) puede solucionar parcialmente el problema, pero persisten limitaciones por la transferencia de coherencia a otros núcleos si no se deuteran los protones no intercambiables. Por este motivo, es aconsejable estudiar los sistemas de masa molecular grande en campo alto, con experimentos de detección de ^{13}C , siendo la detección de los carbonos alifáticos lo que resulta más prometedor para obtener información de alta resolución ya que conservan señales relativamente estrechas (Felli & Brutscher 2009). Por otra parte, como los valores de relajación cruzada homonuclear son más grandes para moléculas cuya rotación es más lenta y la difusión de espín dentro de una red de carbonos enlazados directamente es muy eficiente, los experimentos NOESY ^{13}C - ^{13}C (Bertini et al. 2004b; Bertini et al. 2004c) se presentan como la alternativa para este tipo de sistemas, figura 2.6.

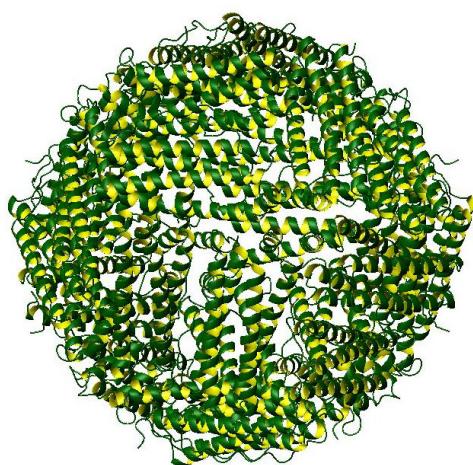


Figura 2.6. Proteína de gran tamaño. Ferritina de rana toro, formada por 24 monómeros (176 aminoácidos por monómero) cuya caracterización en disolución se ha realizado con el uso combinado de varios experimentos NOESY ^{13}C - ^{13}C . Código pdb de la imagen 3RE7 (Bertini et al. 2012).

1.2.2.3. Proteínas y regiones desestructuradas

Ya se ha comentado en el capítulo I que las Proteínas Intrínsecamente Desordenadas (IDP) se caracterizan por carecer de una estructura tridimensional definida (Tomba 2002; Mittag & Forman-Kay 2007). Al contrario de lo que ocurre con las proteínas globulares, las IDPs no se pliegan espontáneamente y mantienen una conformación aleatoria característica de su estado nativo.

Debido al alto grado de desorden y flexibilidad, las proteínas intrínsecamente desordenadas pueden adoptar conformaciones múltiples y variables, haciendo imposible su estudio por otras técnicas como la cristalografía, ya que la preparación de cristales ordenados no es viable. De esta manera, la RMN se posiciona como la herramienta indispensable para la caracterización de dichas proteínas.

Sin embargo, la caracterización biofísica de las IDPs por RMN tiene que superar algunas limitaciones y dificultades debido a sus propias propiedades, que provocan un drástico impacto en los parámetros de RMN, afectando a la dispersión del desplazamiento químico, las constantes de acoplamiento escalar, los acoplamientos dipolares residuales, los NOEs, el intercambio con el solvente, etc. A pesar de esto, los parámetros experimentales que pueden obtenerse nos proporcionan una gran información que resulta indispensable para una mayor comprensión de las reglas que rigen la funcionalidad y dinámica de este tipo de proteínas.

La ausencia de una única estructura tridimensional estable, también puede deberse al efecto de agentes químicos o físicos, que conducen a la desnaturalización total o parcial de una proteína. Por otra parte, hay proteínas, en general globulares, que en su estado nativo presentan regiones o dominios desestructurados como la que se muestra en la figura 2.7.



Figura 2.7. Proteína con región desordenada. Forma completamente reducida de Cox17, código pdb 2RNB (Banci et al. 2008), para la cual se ha combinado el uso de experimentos de detección de ^1H y ^{13}C , lo que ha permitido la asignación secuencial de más del 99% de dicha proteína.

Ya se ha comentado que este desorden conformacional, compartido por todos los sistemas anteriormente mencionados, tiene un impacto importante en la espectroscopía de RMN de protón donde la reducción de la dispersión de los desplazamientos químicos nucleares es muy grande y se pierden las señales de los grupos H_N por el elevado intercambio intrínseco con el solvente. Por ello, en el estudio de los sistemas altamente flexibles, se hace imprescindible el uso de la RMN de detección de ^{13}C , donde la dispersión de los desplazamientos químicos correspondientes a los núcleos de ^{13}C continúa siendo buena (Dyson & Wright 2004), y donde las señales de ^{13}C - ^1H no se ven afectadas por procesos de intercambio, conservando una anchura moderada.

1.2.2.4. Sistemas con intercambio rápido con el solvente

Finalmente, la detección directa de ^{13}C supone una ventaja frente a la detección directa de ^1H , en proteínas que no estando desplegadas, presentan procesos dinámicos importantes de intercambio de los ^1H amídicos H_N con el solvente. Ello hace que los $^1\text{H}_\text{N}$ lleguen a ser inobservables, como el caso de la proteína que se muestra en la figura 2.8 (Arnesano et al. 2001), lo que impide por ejemplo la asignación del esqueleto proteico por métodos tradicionales basados en las correlaciones derivadas de este grupo. La detección de ^{13}C es capaz de superar estos problemas y contribuir a la caracterización de estos sistemas.

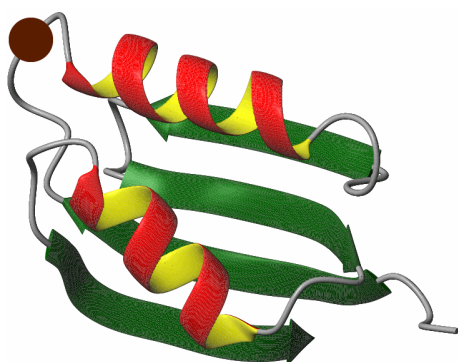


Figura 2.8. Intercambio con el disolvente. La chaperona de cobre Atx1, código pdb 1FD8, en la cual los enlaces metálicos de los bucles en la forma apo de la proteína, que tienen un intercambio muy rápido con el solvente, escapan a los experimentos de detección basados en ^1H .

1.3. Adquisición de datos en RMN

Tradicionalmente la adquisición de datos en RMN se realiza mediante el método lineal convencional, que consiste en adquirir todos y cada uno de los puntos de la FID (Free Induction Decay) en las distintas dimensiones, figura 2.9 (A). Este método de adquisición supone un problema cuando se realizan experimentos de alta dimensionalidad, donde los tiempos de adquisición pueden ser excesivamente largos. De hecho, a menudo se emplean tiempos más cortos de los deseados a costa de reducir drásticamente la sensibilidad y/o la resolución.

Durante los últimos años se han desarrollado nuevos métodos para reducir la duración total de un experimento de RMN modificando la adquisición de datos con aproximaciones, entre otras, de muestreo alternativo (Coggins & Zhou 2006; Marion 2006; Coggins & Zhou 2007; Pannetier et al. 2007). A grandes rasgos, todos estos métodos se basan en la adquisición de un pequeño porcentaje de los puntos del dominio de tiempo de las dimensiones indirectas (Felli & Brutscher 2009). Las posibilidades, figura 2.9, incluyen distribuciones de puntos más ordenadas, como espirales y muestreo radial, y distribuciones completamente desordenadas, como muestreos aleatorios. Sin embargo, la aparición de artefactos cuando se emplean ciertos patrones de muestreo hace que no todas las opciones sean adecuadas.

En este trabajo se ha empleado el muestreo no-uniforme, "non-uniform sampling" (NUS), figura 2.9 (B), en el cual sólo se adquieren en torno al 10-30% de los puntos, en las distintas dimensiones de forma aleatoria (Maciejewski et al. 2006). Recientemente, las evidencias indican que el empleo del muestreo con NUS presenta el nivel más bajo de artefactos (Hoch et al. 2008) en el espectro transformado. De hecho, este esquema de muestreo no causa artefactos importantes. Aun así, los artefactos que pueden generarse, se consiguen distinguir gracias a la diferencia de intensidad que se observa entre ellos y las señales de correlación.

Los espectros adquiridos utilizando este tipo de esquemas de muestreo no pueden procesarse mediante los algoritmos empleados para espectros con muestreo convencional, por lo que requiere implementaciones particulares que permitan procesar este tipo de datos.

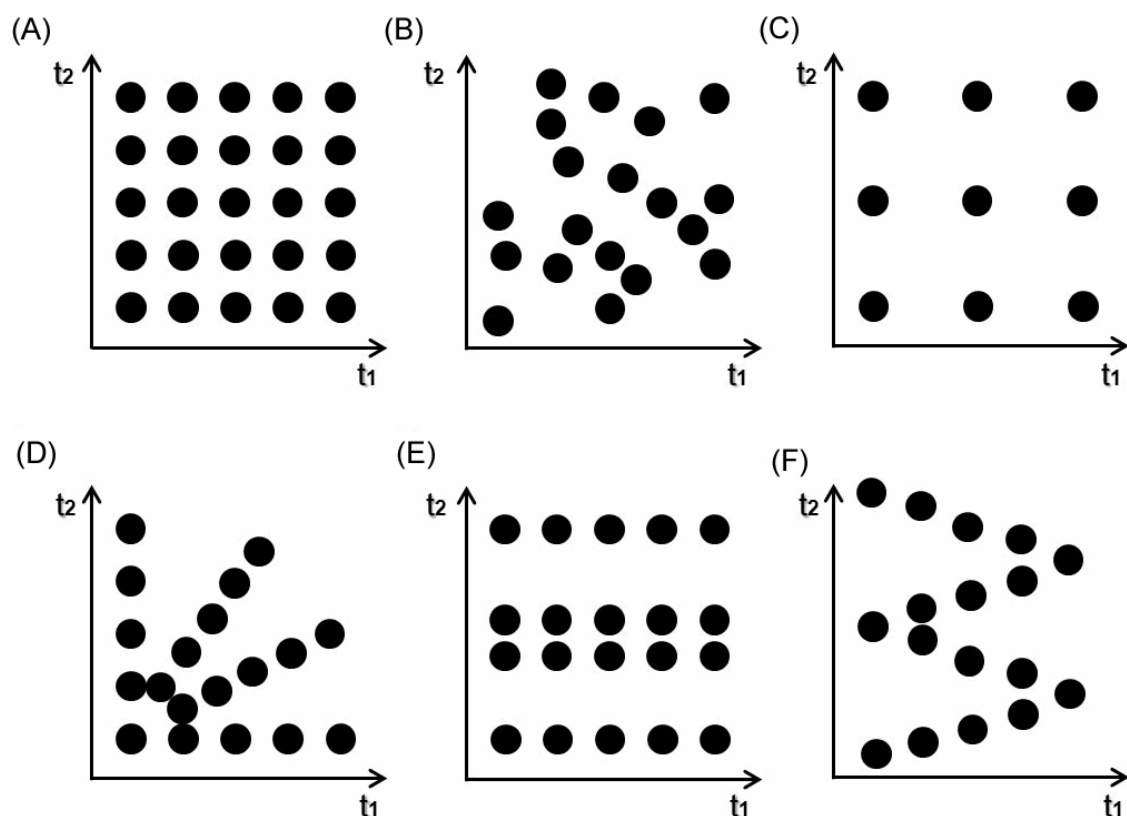


Figura 2.9. Esquemas de muestreo alternativos al muestreo convencional (A), como son el muestreo no-uniforme (B), muestreo lineal reducido (C), muestreo radial (D), muestreo selectivo con Hadamard (E), muestreo ultra-rápido (F).

1.4. Objetivo

El trabajo que se presenta en este capítulo pretende, desde un punto de vista metodológico, explorar la aplicabilidad de la RMN de detección de ^{13}C a las proteínas carentes de estructura, dadas las limitaciones de la aplicación de la RMN de detección de ^1H . Para ello se propone un nuevo método de detección directa de ^{13}C , que complementa a los ya existentes en la literatura para la asignación espectral y que mejora sustancialmente la dispersión de señales. Así, se ha desarrollado una nueva secuencia de pulsos basada en la detección directa del $^{13}\text{C}_\alpha$, optimizada para las IDPs, que sirve para mejorar los métodos convencionales de RMN y aliviar la congestión espectral. El fin último es aportar una herramienta adicional de RMN que facilite la correcta e inambigua asignación espectral en casos en los que las aproximaciones actuales son insuficientes. Para ello se ha trabajado con la α -sinucleína, proteína paradigmática que pertenece a este grupo de proteínas intrínsecamente desestructuradas (Eliezer et al. 2001; Post 2003), que como se ha comentado, están implicadas en gran cantidad de interacciones que condicionan sus funciones biológicas. Este trabajo se ha llevado a cabo durante una estancia en el laboratorio de la Profesora Isabella Caterina Felli del Departamento de Química de la Universidad de Florencia, CERM, Italia. Los resultados están publicados en: Bermel W., Bruix M., Felli I.C., Kumar V., Pierattelli R. and Serrano S. *“Improving the chemical shift dispersion of multidimensional NMR spectra of intrinsically disordered proteins”* J. Biomol. NMR (2013) 55, 231-237.

2. MATERIALES Y MÉTODOS

2.1. Cálculo de los desplazamientos químicos del random coil

La tendencia general que se observa en las IDPs de que los desplazamientos químicos de los núcleos de los aminoácidos tienden a agruparse alrededor de los valores predichos para cada tipo de aminoácido es bien conocida (Tamiola et al. 2010). Estos valores experimentales pueden usarse como base para obtener los desplazamientos del estado random coil (δ_{rc}).

En este trabajo, para el cálculo de los valores de δ del random coil, se han utilizado: (I) los desplazamientos químicos experimentales de la α -sinucleína (BMRB con número de acceso 6968) (Bermel et al. 2006b) modificados con las correcciones, para cada posición, según la naturaleza de los residuos ($i\pm 1$) e ($i\pm 2$) (Wishart et al., 1995; Schwarzingner et al., 2001; De Simone et al. 2009; Kjaergaard and Poulsen, 2011) y (II) los valores determinados por Kjaergaard y colaboradores (Kjaergaard et al. 2011; Kjaergaard & Poulsen 2011) para los átomos de C_α , C_β , C' , N_H , H_α y H_N medidos en una serie de pentapéptidos con las correcciones de secuencia correspondientes.

En un caso general de proteínas sin estructura, estos δ_{rc} calculados pueden ser usados para conocer, con anterioridad a la adquisición de un experimento, las posiciones aproximadas de los picos de cruce en los diferentes espectros de correlación que puedan plantearse a partir de la simulación de los espectros.

2.2. Muestra de RMN empleada

Para la adquisición de los espectros se utilizó una muestra 1.0 mM de la proteína α -sinucleína uniformemente etiquetada en ^{13}C y ^{15}N , disuelta en tampón de fosfato 20 mM a pH 6.5 a 13 °C. Se añadieron EDTA y NaCl hasta alcanzar una concentración final de 0.5 mM y 200 mM respectivamente. La muestra contiene un 10% de D_2O para ajustar el lock. La proteína fue preparada como se describe en la literatura (Huang et al. 2005).

2.3. Experimentos de RMN

Los experimentos se han adquirido en un espectrómetro de RMN Bruker Avance con un campo magnético de 16.4 Teslas que opera a una frecuencia de 700.06 MHz para el ^1H y 176.03 MHz para el ^{13}C , equipado con una sonda criogénica, TXI, optimizada para la detección directa de ^{13}C . Todos los experimentos de RMN, excepto el que se describe en el apartado de resultados, se obtuvieron con secuencias de pulsos descritas previamente en la

literatura. ^1H - ^{15}N HSQC y ^1H - ^{13}C -HSQC (Bodenhausen & D.J. 1980), CON (Bermel et al. 2006b), intra-CAN (Takeuchi et al. 2008; Takeuchi et al. 2010). Los detalles de los experimentos CON con NUS y sin NUS se especifican en el pie de figura de los respectivos espectros. De modo general, los experimentos con NUS se realizaron muestreando aleatoriamente $\frac{1}{4}$ de los puntos del muestreo total convencional. Para el procesado de estos experimentos se ha empleado la Transformada de Fourier Multidimensional (MFT), que resulta muy útil para la reconstrucción de espectros de RMN multidimensionales, concretamente la aproximación “Sparse Multidimensional Fourier Transform” (SMFT). Los detalles experimentales de la secuencia que se ha desarrollado en este trabajo se describen en el apartado de Resultados (3.4).

Previamente al análisis, se calibraron los desplazamientos para el protón, nitrógeno y carbono. Los desplazamientos para el protón y el carbono se referenciaron con respecto a las señales de ^1H y ^{13}C del DSS (2,2-dimetil-2-silapentano-5-sulfonato sódico) a 0.0 y 0.0 ppm, respectivamente. Los desplazamientos de nitrógeno se calibraron indirectamente (Markley et al. 1998).

Todos los datos fueron adquiridos utilizando el programa Bruker TopSpin 1.3 (Bruker Biospin, Karlsruhe, Alemania) y procesados con los programas Bruker TopSpin 1.3 y NMRPipe.

2.4. Análisis y asignación espectral

Para analizar los espectros se utilizaron los programas CARI (Keller 2004) y SPARKY (Goddard 2005). En particular, SPARKY se usó con el fin de identificar las frecuencias de los picos y crear las correspondientes listas y CARI para la identificación de señales y la preparación de diagramas e imágenes. La asignación espectral se ha llevado a cabo a partir de los datos de δ para la α -sinucleína, código BMRB 6968.

3. RESULTADOS

3.1. Limitación en la detección directa de ^1H . Necesidad de otras aproximaciones

Para analizar las limitaciones de los experimentos de detección de ^1H cuando se trabaja en sistemas carentes de estructura, hemos utilizado los experimentos clásicos 2D de detección de ^1H , que son la base de la batería de experimentos 3D-RMN necesaria para la asignación espectral de proteínas. La proteína tomada como modelo en este estudio ha sido la α -sinucleína. Esta proteína tiene una gran importancia biológica, por estar relacionada con diversas enfermedades degenerativas, entre ellas el Parkinson (Eriksen et al. 2003). Está compuesta por 140 residuos de los cuales 5 son prolinas y 17 glicinas. Se trata de una proteína, que al carecer de estructura tridimensional, manifiesta plasticidad conformacional y que pertenece al grupo de las proteínas intrínsecamente desordenadas.

Como ya se ha comentado en la Introducción, los espectros de protón de estas proteínas se caracterizan por la poca dispersión de las señales. Para visualizar este efecto se registraron los espectros ^1H - ^{15}N y ^1H - ^{13}C HSQC que se muestran en las figuras 2.10 y 2.11, respectivamente. Como puede observarse, los valores de los δ tienden a converger en determinadas regiones del espectro.

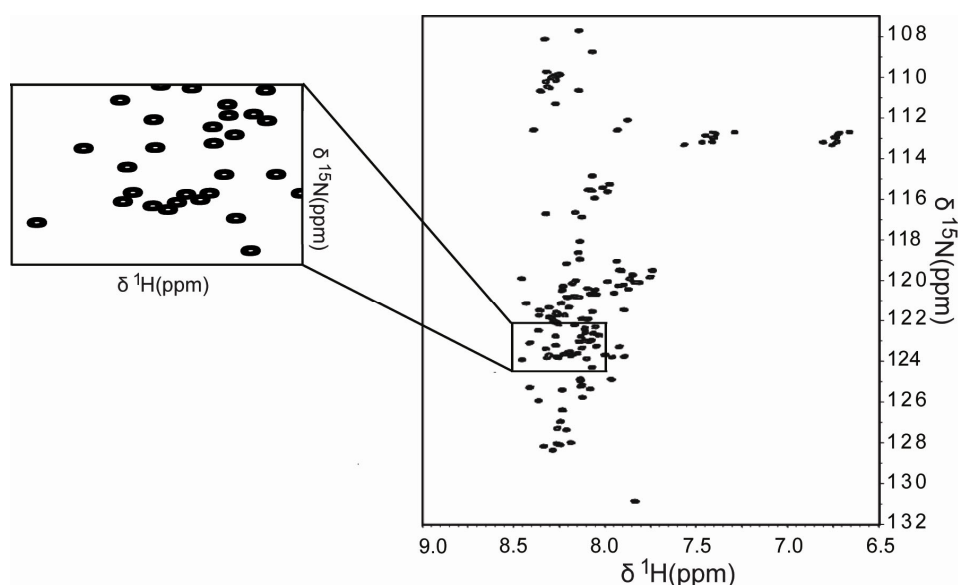


Figura 2.10. Experimento ^1H - ^{15}N HSQC, de una muestra de la proteína α -sinucleína marcada en ^{13}C y ^{15}N a una concentración de 1.0 mM, que muestra las correlaciones $^1\text{H}_\text{N}$ - $^{15}\text{N}_\text{H}$ intrarresiduales. El experimento ha sido adquirido en un espectrómetro Bruker Avance de 16.4 Teslas a una temperatura de 285.5 K.

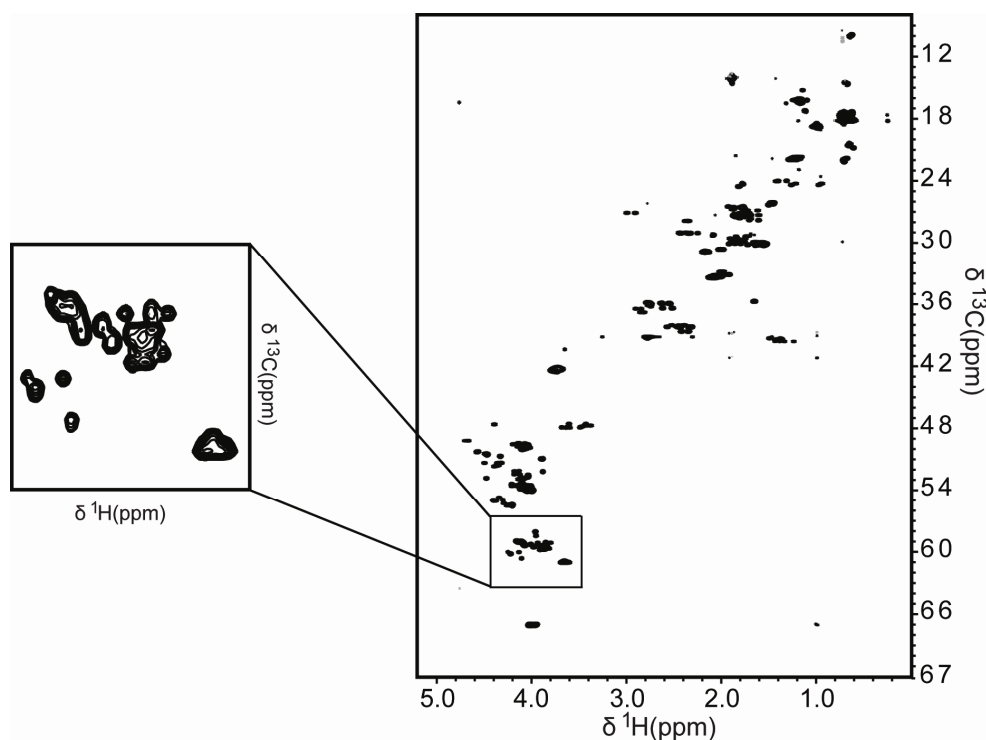


Figura 2.11. Experimento ^1H - ^{13}C HSQC dónde se observan las correlaciones ^1H - ^{13}C intrarresiduales de la proteína α -sinucleína 1.0 mM, a una temperatura de 285.5 K, adquirido en un espectrómetro Bruker Avance de 16.4 Teslas que opera a una frecuencia de 700 MHz para el ^1H .

Por otra parte, dado que las anchuras espectrales típicas que deben cubrirse varían de un núcleo a otro, la resolución digital que se consigue en condiciones normales es diferente para las dimensiones de ^{15}N y de ^{13}C . La zona a cubrir para las correlaciones directas ^1H - ^{15}N amida es de unos 35 ppm aproximadamente, mientras que para los ^{13}C alifáticos es de unos 70 ppm. Si comparamos regiones de tamaños idénticos en hercios, figura 2.12, tanto para la dimensión directa como para la indirecta (2.5 ppm para la dimensión directa de ^1H , 25 ppm para la dimensión indirecta de ^{15}N y 10 ppm para el ^{13}C) de los experimentos ^1H - ^{13}C y ^1H - ^{15}N HSQC, comprobamos que para las señales de ^{13}C existe un solapamiento mayor dada la resolución digital menor que suele obtenerse en los experimentos convencionales.

Es por ello que puede resultar de interés mejorar las aproximaciones que se basan en la detección de ^{13}C para proteínas sin estructura, en las que se pueda aliviar el problema del solapamiento de señales.

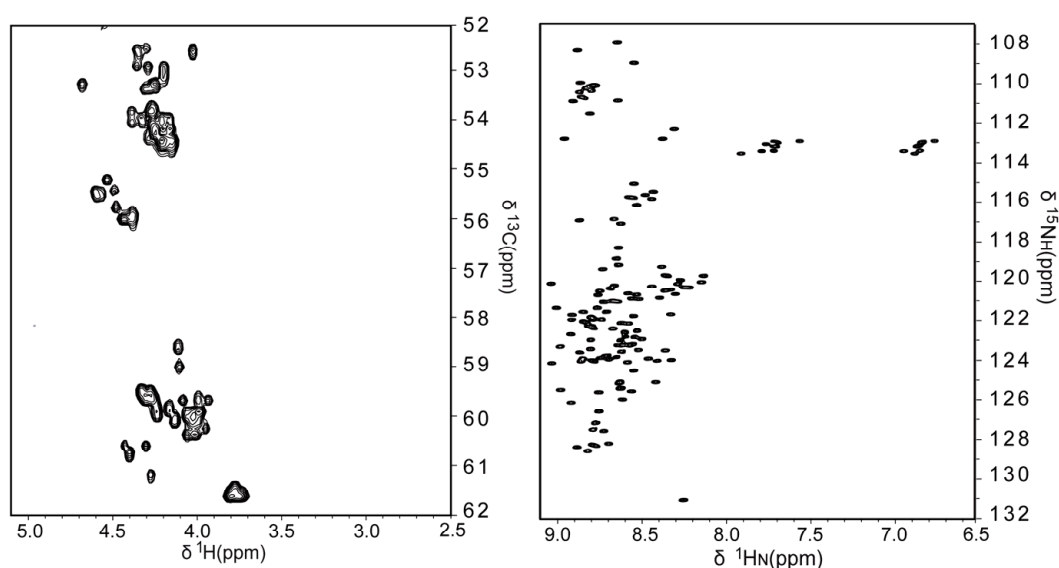


Figura 2.12. A la izquierda, región del espectro ^1H - ^{13}C HSQC y, a la derecha, espectro ^1H - ^{15}N HSQC. Ambos espectros están representados con una anchura espectral para la dimensión directa de ^1H de 2.5 ppm y para la dimensión indirecta 10 ppm para el ^{13}C y 25 ppm para el ^{15}N . Estos valores en ppm equivalen a 1750 Hz para ambas dimensiones.

3.2. Exploración de las zonas espectrales de interés en detección de ^{13}C para proteínas sin estructura

Con el fin de hacer un planteamiento lo más general posible de las zonas donde se encontrarán los picos de cruce en los distintos experimentos con detección de ^{13}C -RMN para proteínas sin estructura, se han representado los δ_{rc} en función de las diferentes correlaciones: $\text{C}'\text{-C}_\alpha$, $\text{C}'\text{-N}_\text{H}$ y $\text{C}_\alpha\text{-N}_\text{H}$, tanto intrarresiduales como interresiduales para la sinucleína. Las figuras 2.13 y 2.14 muestran los diagramas calculados empleando los dos grupos de valores de δ_{rc} (según se describe en Materiales y Métodos). Los paneles obtenidos con las dos bases de datos son muy similares a excepción de los picos de las señales correspondientes a las Pro. Este hecho se debe a que para estos residuos, no están determinados los parámetros de δ_{rc} para el ^{15}N en las series de pentapéptidos utilizados como referencia (Bermel et al. 2006b; Kjaergaard et al. 2011; Kjaergaard & Poulsen 2011). Por este motivo, es más completa la aproximación basada en los datos experimentales publicados de la α -sinucleína.

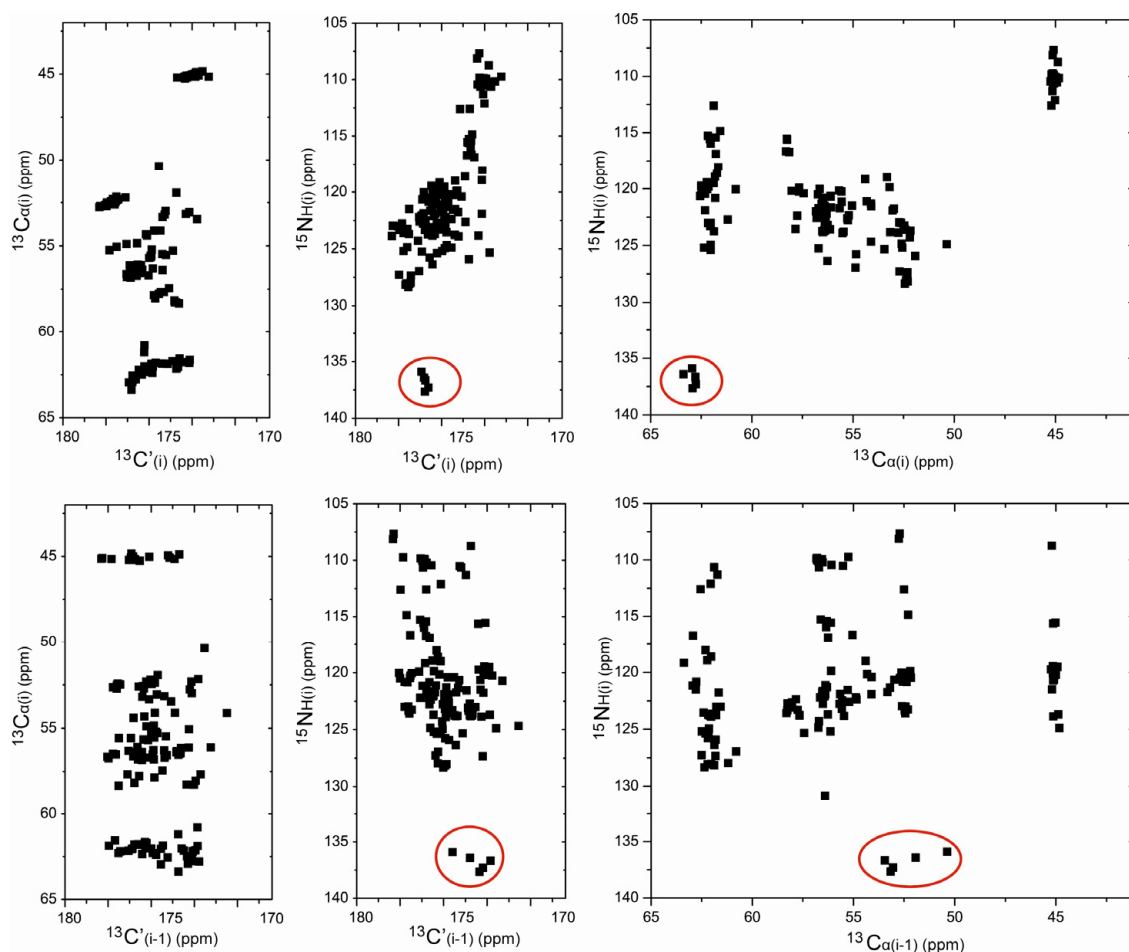


Figura 2.13. Gráficas de correlación $\text{C}'\text{-C}_\alpha$, $\text{C}'\text{-N}_\text{H}$ y $\text{C}_\alpha\text{-N}_\text{H}$ obtenidas usando los desplazamientos químicos δ_{rc} a partir de los de la α -sinucleína (BMRB 6968) (Bermel et al., 2006a). Los paneles superiores, muestran las correlaciones intrarresiduales y los inferiores las correlaciones interresiduales. La figura ha sido previamente publicada en el artículo (Bermel et al. 2013). En el círculo rojo se muestran las correlaciones $^{13}\text{C}\text{-}^{15}\text{N}$ correspondientes a los residuos Pro.

De las figuras 2.13 y 2.14, podemos deducir en primer lugar el porqué el experimento CACO (correlación intrarresidual entre el C_α y el C') que es altamente sensible y que se aplica con mucho éxito en proteínas globulares (Takeuchi et al. 2008) no tiene la misma aplicabilidad en IDPs. Esto es debido a que, en el caso de las proteínas globulares, existen contribuciones adicionales a la naturaleza de cada residuo que influyen en la mayor dispersión de las señales en el espectro. En segundo lugar, es relevante el hecho de que las señales de los C_α tienen mayor dispersión que otros átomos, de casi unas 20 ppm, ya que sus valores más bajos, correspondientes a Gly aparecen en torno a 45 ppm y los más altos en torno a 63 para las Pro. En tercer lugar, podemos ver que en la representación anterior, se aprecia claramente una mayor dispersión de las señales cuando reproducen las

correlaciones interresiduales (paneles inferiores) frente a las intrarresiduales (paneles superiores). Este hecho resulta sumamente interesante y abre la puerta para el desarrollo de nuevas secuencias de pulsos, como la que se desarrolla en este trabajo.

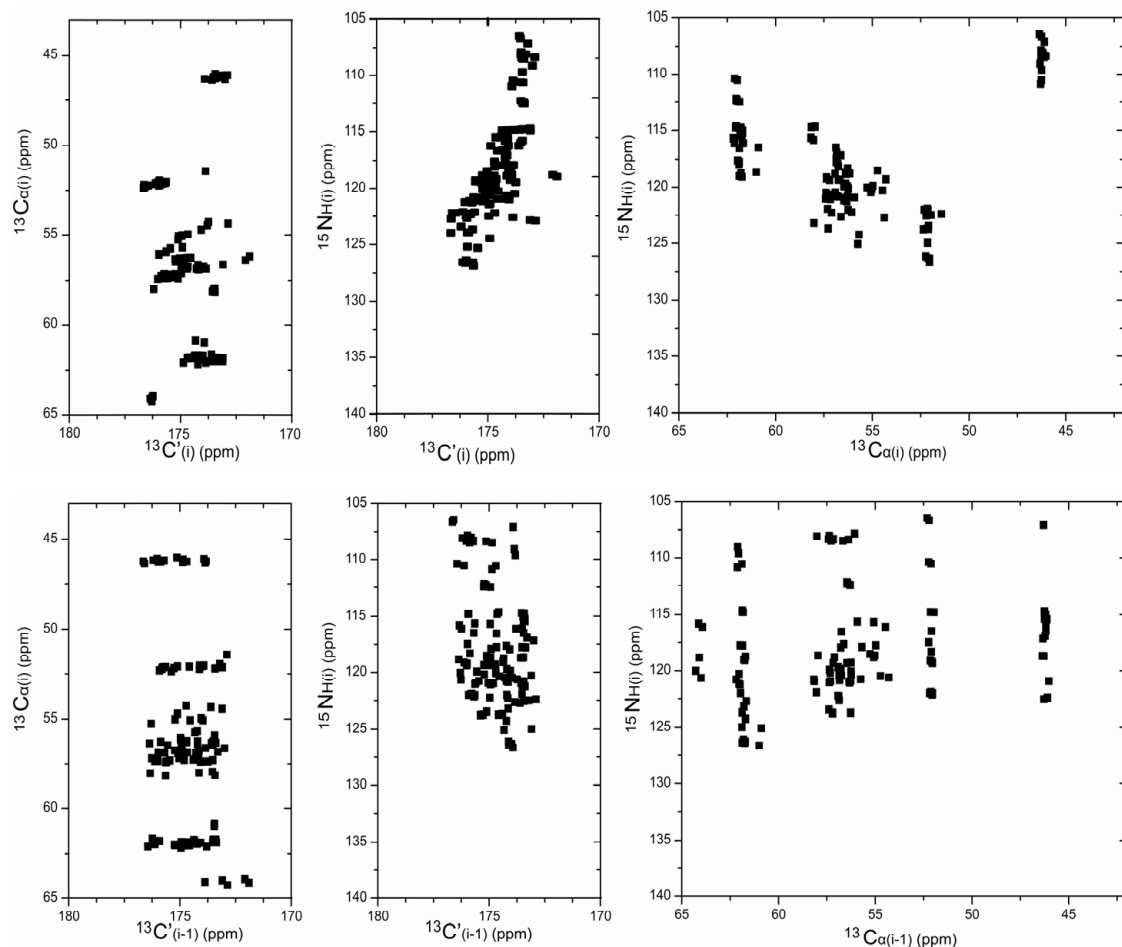


Figura 2.14. Gráficas de correlación entre $C'-C_{\alpha}$, $C'-N_H$ y $C_{\alpha}-N_H$ obtenidas usando los desplazamientos químicos δ_{rc} teóricos calculados (Kjaergaard & Poulsen 2011). Los paneles superiores, muestran las correlaciones intrarresiduales y los inferiores las correlaciones interresiduales.

3.3. Mejoras en la dispersión de señales con la detección de $^{13}\text{C}'$ y NUS

Para mejorar estos problemas de solapamiento se ha propuesto una metodología de asignación secuencial (Bermel et al. 2006c) basada en la detección de $^{13}\text{C}'$ de igual manera que existe para la detección de ^1H . En la figura 2.15 se comparan los dos espectros 2D, que podrían tomarse como la huella dactilar de la α -sinucleína, ^1H - ^{15}N HSQC, que emplea la detección de ^1H , y CON (experimento que relaciona el carbono carbonílico con el nitrógeno del siguiente aminoácido, unido directamente del esqueleto polipeptídico), que emplea la detección de $^{13}\text{C}'$, ambos obtenidos en condiciones similares. Puede verse que la detección del carbono carbonílico provoca una mejora importante en la dispersión de los picos. De esta manera, se resuelven aproximadamente el 90% de las correlaciones predichas, facilitando así la asignación secuencial. Además, en el espectro 2D CON los residuos de Pro pueden ser identificados fácilmente demostrando como este último experimento, nos proporciona información importante para el proceso de asignación.

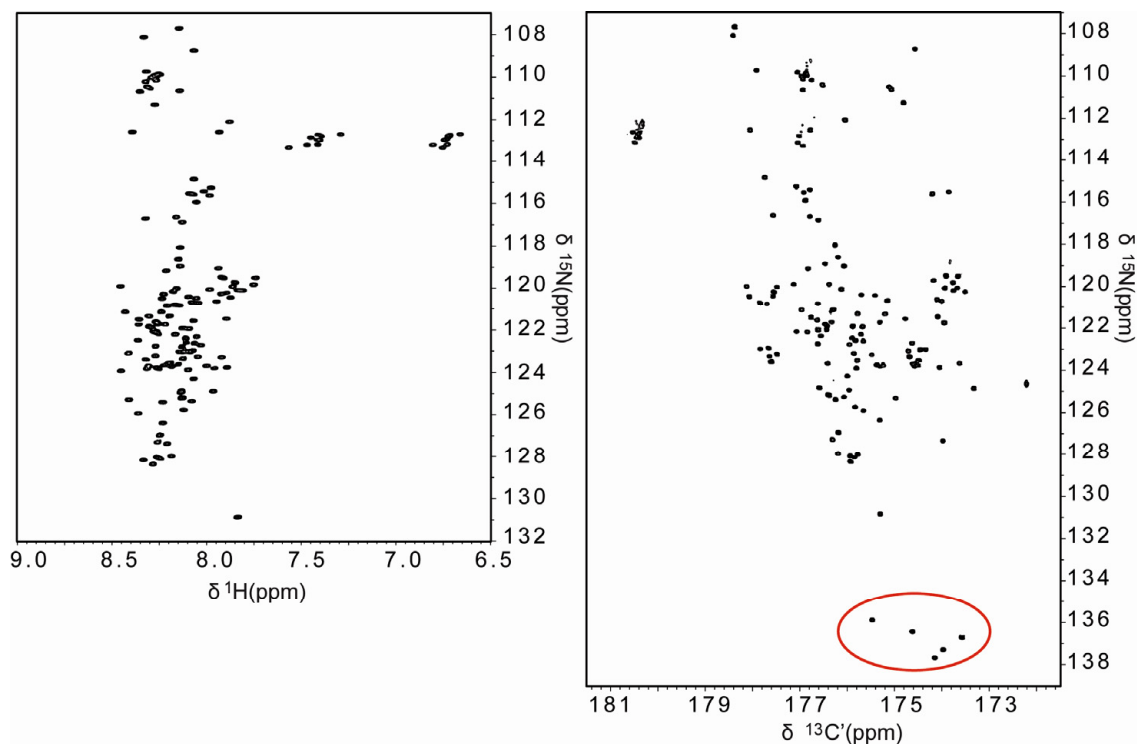


Figura 2.15. Experimentos adquiridos de una muestra 1.0 mM de α -sinucleína en un espectrómetro Bruker Avance de 16.4 T. A la izquierda experimento ^1H - ^{15}N HSQC en el que el tiempo total de adquisición fue de 10 minutos y a la derecha experimento CON, adquirido en 40 minutos. Ambos espectros están representados con la misma anchura espectral de 1750 Hz en la dimensión directa. En círculo rojo se muestran las correlaciones correspondientes a los residuos Pro.

Para poner de manifiesto otra de las ventajas que tiene la RMN de detección directa de ^{13}C , respecto a la detección de ^1H , se llevaron a cabo una serie de experimentos ^1H - ^{15}N HSQC y CON a diferentes temperaturas. Evidentemente se mantiene la mejor dispersión de señales en la detección de ^{13}C que en la detección de ^1H , pero además se evita, como se muestra en la figura 2.16, la pérdida de señales por el intercambio de los $^1\text{H}_\text{N}$ con el solvente al aumentar la temperatura (Englander et al. 1972; Mulder et al. 1996).

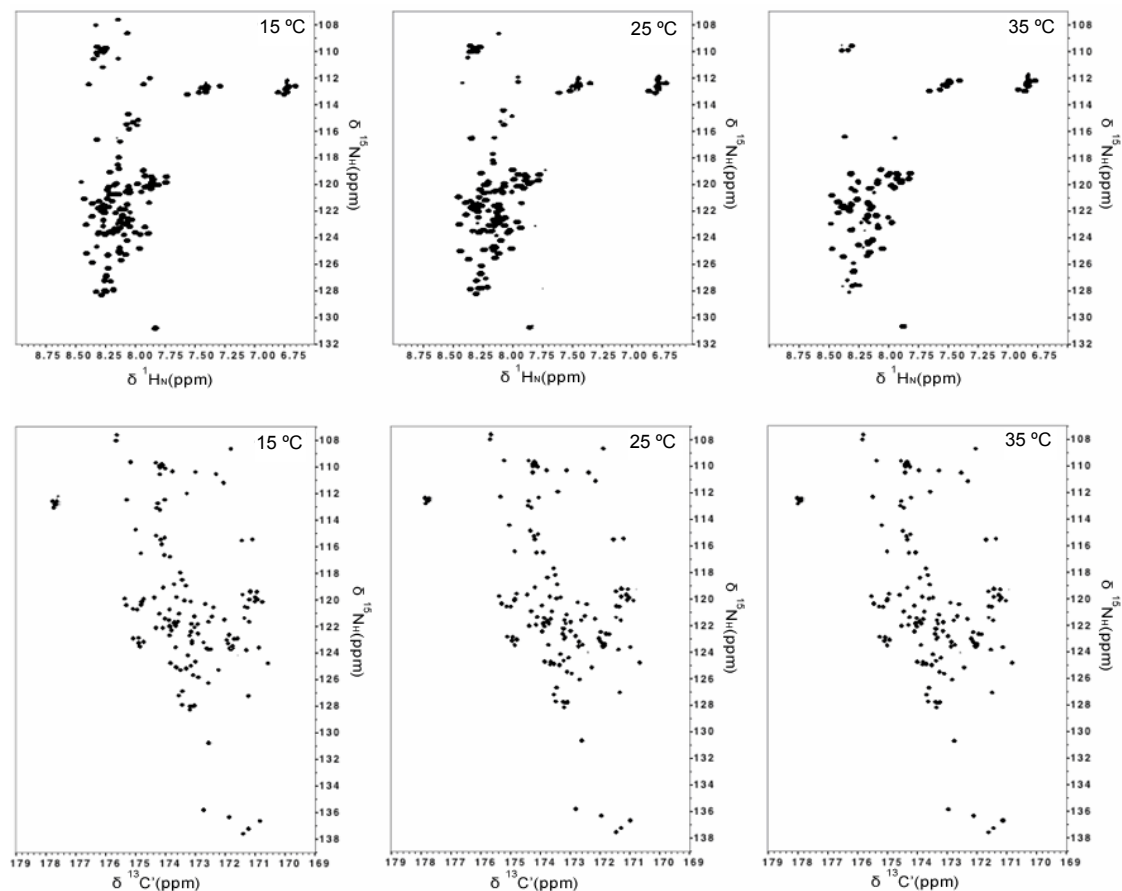


Figura 2.16. Experimentos adquiridos de una muestra 1.0 mM de α -sinucleína en un espectrómetro Bruker Avance de 16.4 T a diferentes temperaturas: 15 °C, 25 °C y 35 °C. En la parte superior experimentos ^1H - ^{15}N HSQC y en la inferior experimentos CON (las escalas de las dimensiones de ^1H y ^{13}C son las mismas en hercios).

Mientras que en los espectros CON el número de señales permanece invariable (se detectan 137 señales individuales de un total de 140 posibles), en el caso de ^1H - ^{15}N HSQC las cosas son bien distintas. A la temperatura de 15 °C, dejando aparte las señales que corresponden a las cadenas laterales, se pueden diferenciar 119 señales de correlación. Este número desciende a 104 cuando el espectro se registra a la temperatura de 25 °C (reducción 12.6 %) y hasta 67 a 35 °C (reducción 43.7 %).

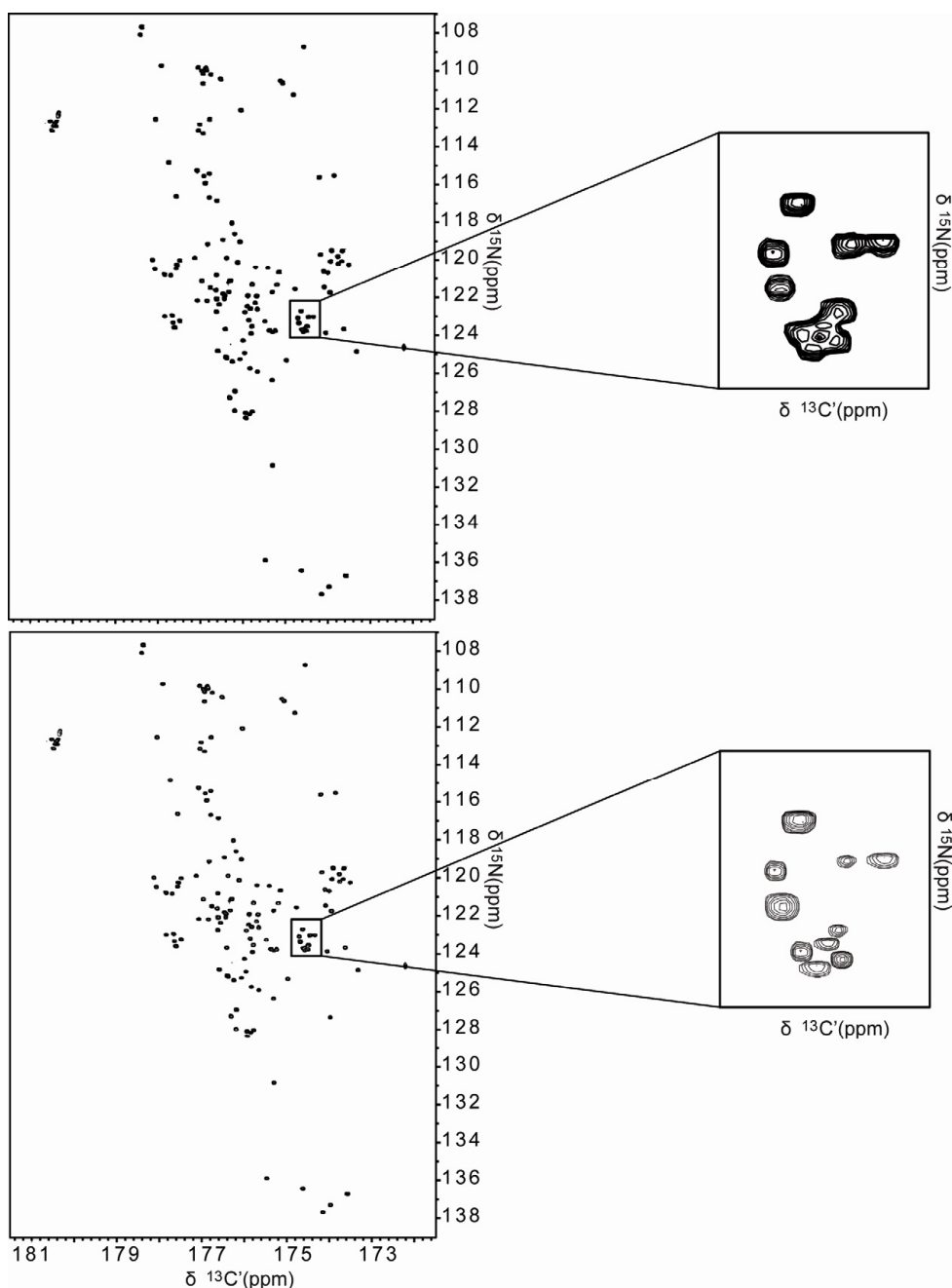


Figura 2.17. Espectros 2D CON adquiridos empleado el método convencional, en la parte superior, y NUS, en la parte inferior. Ambos experimentos se realizaron en un espectrómetro Bruker Avance de 16.4 T en una muestra de 1.0 mM de α -sinucleína. El tiempo total de adquisición de ambos experimentos fue el mismo, 40 minutos, y se usó un barrido por incremento. Las zonas recuadradas muestran la mejora obtenida en el espectro que implementa NUS donde la anchura de las líneas es menor gracias al incremento en el tiempo de evolución del ^{15}N .

Por otra parte, el empleo de NUS es especialmente importante dado que reduce el tiempo experimental o permite aumentar la resolución en la dimensión indirecta. La mejora de la resolución producida por NUS se puede observar en la figura 2.17, en la cual el espectro

2D CON de la α -sinucleína, fue adquirido empleando el muestreo no uniforme y el convencional. Ambos experimentos fueron realizados empleando el mismo tiempo de adquisición pero la resolución, en el caso de NUS, es más grande puesto que el tiempo máximo de evolución del ^{15}N se duplica (de 205 a 410 milisegundos) a pesar de mantener el mismo tiempo total. Ese aumento de resolución pueden observarse claramente en la ampliación insertada de la figura 2.17.

Para demostrar que la reducción del tiempo experimental empleado en los espectros adquiridos con NUS no disminuye la calidad espectral, se adquirió un experimento 2D CON adicional en la mitad de tiempo y se comparó con el realizado mediante muestreo convencional, figura 2.18. Ambos espectros son similares aunque en el adquirido con NUS en la mitad de tiempo se observa un débil decremento en la intensidad de las señales. Además, en el espectro muestreado con NUS, se encontraron un par de picos adicionales que están indicados con flechas en la figura 2.18, que no aparecen en el espectro convencional. Estos picos corresponden a artefactos que pueden ser eliminados posteriormente en el post-procesado con un protocolo apropiado. Por tanto, este enfoque parece ser adecuado para ser implementado en experimentos más complejos como el que se describe más adelante.

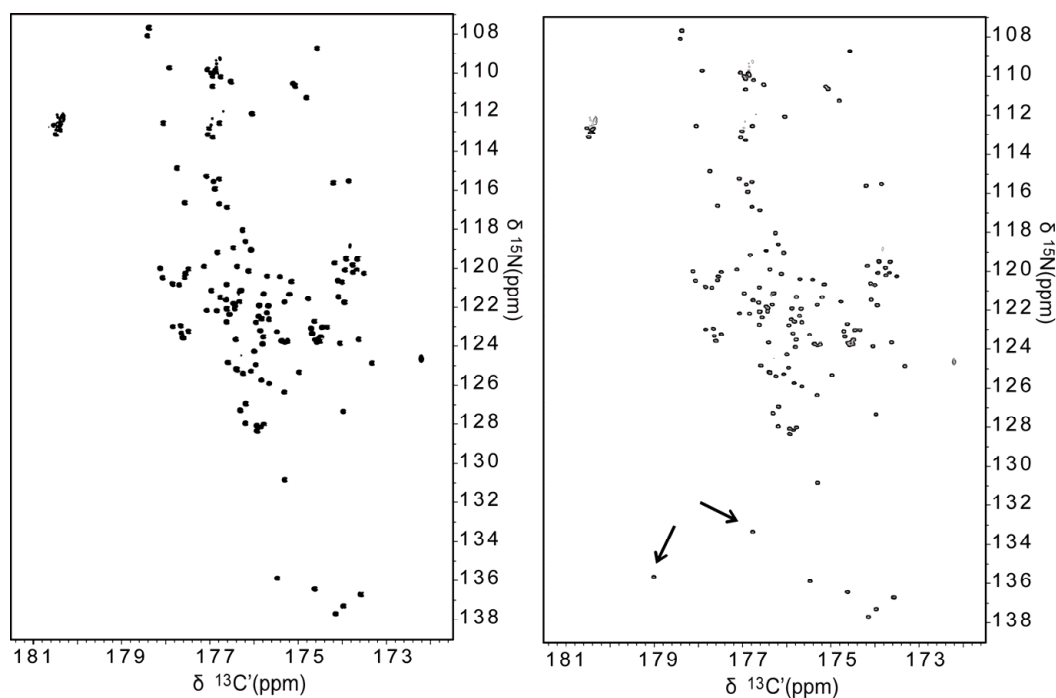


Figura 2.18. Espectros 2D CON. Ambos experimentos fueron adquiridos en un espectrómetro Bruker Avance de 16.4 T en una muestra de 1.0 mM de α -sinucleína. A la izquierda experimento muestreado de forma convencional y a la derecha con NUS. La duración del experimento convencional fue de 40 minutos mientras que el experimento con NUS se realizó en 20 minutos usando 1 barrido por incremento. Las flechas en el panel de la derecha indican los dos artefactos más fuertes que aparecen en el espectro muestreado con NUS.

3.4. Experimento de RMN inter-CAN

Para cumplir con el objetivo de diseñar experimentos útiles para el estudio de las IDPs, nos hemos centrado en las correlaciones $^{13}\text{C}_\alpha$ - ^{15}N interresiduales, que presentan mayor dispersión que las intrarresiduales en los “espectros simulados” (mostrados en el apartado 3.2). Para ello se ha diseñado un experimento de RMN que permita correlacionar los $\text{C}_\alpha(i)$ y $\text{N}_\text{H}(i+1)$ de cada uno de los sistemas de espín para conseguir la asignación secuencial de manera rápida y fiable, incluso en presencia de las dificultades que supone el estudio de las proteínas desplegadas.

En el experimento bidimensional inter-CAN que se ha diseñado, la magnetización sigue el camino indicado en la figura 2.19.

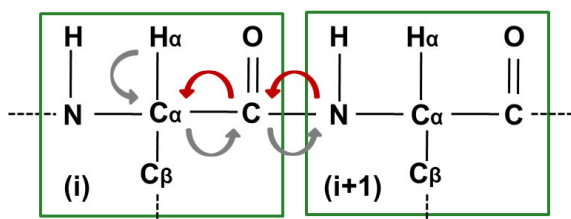


Figura 2.19. Camino de la transferencia de magnetización en el experimento bidimensional inter-CAN que se inicia en el H_α del residuo (i), llega hasta el N_H del (i+1) y vuelve al C_α del residuo (i). Las flechas grises representan el camino de la transferencia de polarización de ida y las flechas rojas el retorno de la misma.

Independientemente del protocolo usado para la transferencia de coherencia, como se ha mencionado en la introducción, los espines de los protones deben ser los primeros en ser excitados para mejorar la relación señal/ruido ya que ésta depende tanto de la constante giromagnética del núcleo observado al final de la secuencia como de la constante del núcleo excitado al inicio de la misma. Así pues, aquí se comienza en el H_α del residuo (i), pasa a través del C_α al C' del mismo residuo (i), para llegar hasta el N_H del residuo siguiente (i+1) donde se recogen las frecuencias de resonancia del N_H , que darán lugar a la dimensión indirecta. A continuación la magnetización retorna al C' y por último al C_α , cuyas resonancias dan lugar a la dimensión directa.

La figura 2.20 muestra el esquema de la secuencia de pulsos inter-CAN. Una vez se han excitado los núcleos de ^1H , la secuencia aprovecha el acoplamiento del enlace $J_{\text{C}\alpha\text{H}\alpha}$ (140 Hz) para transferir la magnetización a los heteronúcleos unidos directamente al protón. Después del primer pulso de 90° , la magnetización es transferida al $^{13}\text{C}_\alpha$ a través de la evolución con $^1J_{\text{H}\alpha\text{C}\alpha}$, y luego al $^{13}\text{C}'$, a través de la evolución $^1J_{\text{C}\alpha\text{C}'}$ llegando finalmente al $^{15}\text{N}_\text{H}$ a través de la evolución con $^1J_{\text{C}'\text{N}_\text{H}}$. En este punto se recogen las frecuencias de resonancia del ^{15}N para crear la dimensión indirecta. A continuación, se recogen las resonancias $^{13}\text{C}_\alpha$ de la dimensión directa producidas por el retorno de la magnetización al $^{13}\text{C}'$ y luego al $^{13}\text{C}_\alpha$, a través de las

evoluciones con $^1J_{\text{C'NH}}$ y $^1J_{\text{CaC'}}$, respectivamente. Al final de la secuencia, se emplean dos bloques IPAP (Andersson et al. 1998; Ottiger et al. 1998; Duma et al. 2003) concatenados en un solo bloque más corto, DIPAP (Duma et al. 2003; Bermel et al. 2005a), para lograr el desacoplamiento homonuclear del $^{13}\text{C}_\alpha$.

En el canal selectivo del ^{13}C se han empleado pulsos de $\pi/2$ y π con forma Q5 (o Q5 de tiempo inverso) y Q3 (Emsley and Bodenhausen, 1992) con una duración de 300 y 220 μs , respectivamente, a excepción de los pulsos de π empleados en la región de los C_α (Q3 de 860 μs) y el pulso adiabático de π que se ha empleado para invertir tanto C' como C_α (con un pulso suave de 500 μs , homogéneo al 25%, un barrido de 80 kHz y con una intensidad de campo para la radio-frecuencia de 11.3 kHz (Boehlen and Bodenhausen, 1993)). En el canal selectivo de ^{13}C , se han dado pulsos selectivos en el centro de cada región para C_{all} , C_α , y C' y el pulso adiabático se ha ajustado para cubrir toda la región de ^{13}C . El desacoplamiento de ^1H y ^{15}N se ha realizado mediante las secuencias WALTZ-16 (Shaka et al. 1983) (1.7 kHz) y GARP-4 (Shaka et al. 1985) (1.0 kHz), respectivamente. Todos los gradientes empleados han tenido una duración de 1 ms.

En los experimentos que emplean el bloque DIPAP (Bermel et al., 2005a; Bermel et al., 2006b), para cada incremento en la dimensión indirecta, se han recogido cuatro FIDs que han sido almacenadas por separado, las matrices resultantes se han sumado y restado por pares para separar las cuatro componentes del multiplete, en ese momento desplazadas por $\pm J_{\text{C}'\text{C}_\alpha}/2$ Hz y $\pm J_{\text{C}_\alpha\text{C}_\beta}/2$ Hz y sumadas de nuevo para obtener un singlete, figura 2.2 (C). En el caso de las Gly, que tienen un sólo acoplamiento, al carecer de C_β , para obtener las señales de manera correcta debe transformarse el espectro empleando un bloque IPAP sencillo. Se ha empleado también la aproximación de selección de estados de espín para desacoplar el ^{15}N del C' , de manera que no se necesita una irradiación de radiofrecuencia sobre el ^{15}N para desacoplarlo, lo cual normalmente limita la duración del tiempo de adquisición y permite aumentar la adquisición de la FID sin las limitaciones impuestas por el desacoplamiento del mismo (Bermel et al. 2009).

Los duración de los tiempos de espera usados ha sido: $\Delta_1 = 7.2$ ms, $\Delta = 9$ ms, $\Delta_2 = 2.2$ ms, $\Delta_3 = 3.6$ ms, $\Delta' = 24.8$ ms, $\Delta'' = 26.6$ ms y $\varepsilon = t_1(0)+p\text{C}180$. Los ciclos de fases empleados han sido: $\Phi_1 = 4(x), 4(-x)$; $\Phi_2 = 2(x), 2(-x)$; $\Phi_3 = x, -x$; $\Phi_4 = 8(x), 8(-x)$; $\Phi_{\text{DIPAP}}(\text{IPIP}) = x$; $\Phi_{\text{DIPAP}}(\text{APIP}) = -y$; $\Phi_{\text{DIPAP}}(\text{IPAP}) = -y$; $\Phi_{\text{DIPAP}}(\text{APAP}) = -x$; $\Phi_{\text{rec}} = x, 2(-x), x, -x, 2(x), -x$. Para la evolución de los desplazamientos químicos de los protones se ha empleado “constant time” (Santoro & King 1992) y para los nitrógenos “semi constant time” (Aitio & Permi 2000) lo que nos permite obtener la máxima resolución espectral. La detección de cuadratura en la dimensión indirecta, F1, se obtuvo por incrementos Φ_2 empleando el método TPPI (States et al. 1982; Marion et al. 1989).



Estudio de interacciones de proteínas por RMN

La figura 2.21 muestra el espectro experimental obtenido para la muestra de α -sinucleína empleando la secuencia inter-CAN donde la implementación del bloque DIPAP, utilizado al final de la secuencia de pulsos, permite obtener una única señal para cada una de las correlaciones. Todas las señales del espectro presentan el mismo signo, excepto las señales de las glicinas que aparecen a la derecha del espectro y tienen distinto signo porque se ha empleado un único bloque IPAP para su transformación.

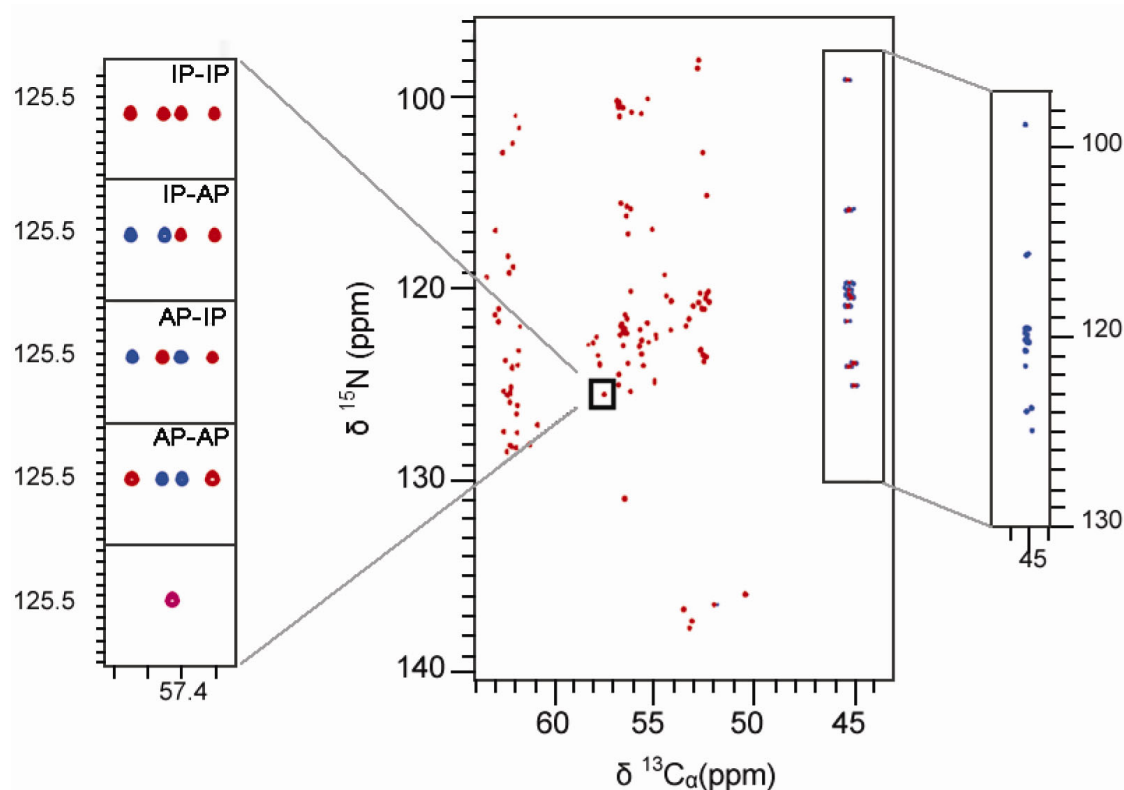


Figura 2.21. Espectro inter-CAN de la ^{13}C , ^{15}N α -sinucleína humana (1 mM en 20 mM tampón fosfato, pH 6.5, 0.5 mM EDTA, 200 mM NaCl) adquirido en un espectrómetro Bruker Avance de 16.4 T. A la izquierda se muestran las distintas posibilidades para un pico arbitrario de los cuatro sub-espectros inter-CAN antes de combinarlos para dar lugar al desacoplamiento homonuclear “virtual” y su combinación para obtener el singlete. En el recuadro de la derecha, región donde se encuentran las resonancias para los residuos de Gly, el espectro se ha procesado utilizando un bloque IPAP sencillo para eliminar únicamente el acoplamiento $\text{C}_\alpha\text{-C}'$. La figura ha sido adaptada del artículo (Bermel et al. 2013).

4. DISCUSIÓN

En este trabajo se han analizado los pros y los contras de las diferentes aproximaciones de detección y condiciones experimentales, y hemos planteado un estudio exhaustivo de la dispersión de señales de correlación ^{13}C - ^{15}N en opciones inter e intrarresidual. En este contexto y después de examinar las secuencias de RMN disponibles en la literatura para la asignación de proteínas sin estructura (Bertini et al. 2011), parece evidente que todavía es necesario el desarrollo de nuevas estrategias capaces de solventar los problemas existentes, expuestos en las secciones anteriores.

Además de la problemática ya descrita anteriormente, el estudio de proteínas o regiones sin estructura por RMN tiene otros problemas asociados que conviene tener en consideración: (I) la estabilidad limitada de las muestras por diversos factores, entre ellos la propensión a la proteólisis, (II) las bajas concentraciones alcanzables en el tubo de RMN para impedir o minimizar la tendencia a agregar, (III) heterogeneidades en las intensidades de los picos de los espectros debidas a diferentes regímenes dinámicos y (IV) repeticiones y degeneración de secuencia de las proteínas. Todo ello implica que, para avanzar en el estudio de estos sistemas, y obtener información estructural y dinámica, se tengan que proponer aproximaciones integradoras no sólo entre diferentes técnicas experimentales si no también con el análisis combinado de diferentes observables de RMN. Así, el análisis conjunto de diferentes observables de RMN puede contribuir a entender detalles importantes relacionados con la función de estas proteínas.

Desde el punto de vista experimental, cualquier diseño de experimentos multidimensionales que pueda hacerse en este campo debe tratar de solventar los problemas mencionados. Las nuevas propuestas deben proporcionar secuencias de pulsos altamente sensibles para obtener información de los picos de diferente intensidad, diferente intercambio con el solvente, baja concentración, etc. Por otra parte, para resolver ambigüedades debidas a la superposición de señales es conveniente conseguir una buena resolución, lo que requiere largos tiempos de adquisición en los experimentos multidimensionales. Sin embargo, esto no es adecuado para trabajar con muestras que tienen tendencia a agregar o que son poco estables en el tiempo. Por ello, independientemente de la aproximación empleada, las técnicas de detección rápida y muestreo no uniforme son obligadas para poder hacer estudios en un tiempo razonable y mantener la integridad de la muestra. Además, las mejoras instrumentales permiten plantearse hoy en día el diseño de experimentos diferentes y complementarios a los clásicos, con detección directa en núcleos menos sensibles que el protón, que permiten aprovechar ventajas importantes como la mayor dispersión espectral y que todas ellas

combinadas permitan la caracterización a nivel atómico de los sistemas biológicos altamente flexibles.

En este contexto, una de las aproximaciones que está dando mejores resultados es la RMN de detección de $^{13}\text{C}'$, tanto en RMN en estado líquido como en estado sólido (Ishii et al. 2001). Desde hace ya unos años se han descrito baterías de experimentos multidimensionales 2D CACO-IPAP (Bermel et al. 2005a), 3D CBCACO-IPAP (Bermel et al. 2005a), 3D CANCO (Bertini et al. 2004a; Bermel et al. 2005b), diversos TOCSYs (Eletsky et al. 2003; Bermel et al. 2005a; Murata et al. 2005; Arinaitwe et al. 2012), etc. para llevar a cabo la asignación y estudio de estas proteínas (Bermel et al. 2006a) de una forma similar a lo que se hace para proteínas globulares con detección de ^1H (Wüthrich 1986), utilizando en la mayoría de los casos métodos que aceleran la adquisición y la resolución. Recientemente se ha publicado la implementación de experimentos basados en la detección directa de $^{13}\text{C}'$ del tipo 4D HCBCACON, HCCCON, HCBCANCO, 4/5D HNCACON y HNCANCO Y 3/4D HCANCAO específicamente diseñados para la identificación de sistemas de spin y la asignación secuencial con buenos resultados (Bermel et al. 2012).

4.1. Nuevas propuestas para el estudio por RMN de IDPs. La secuencia inter-CAN

Si bien es cierto, y se ha comentado que las proteínas intrínsecamente desestructuradas pueden asemejarse a los estados random coil, la realidad desde el punto de vista de los observables de RMN no es tan simple. El examen que se ha hecho en este capítulo, de las representaciones de las señales de correlación observables para la α -sinucleína en distintos experimentos de detección de ^{13}C a partir de los valores calculados de δ random coil, permite sacar conclusiones difíciles de obtener a partir de otros parámetros. Es conocido que los desplazamientos químicos son altamente sensibles al entorno químico del núcleo (Wishart et al. 1995) y que estos entornos se promedian en sistemas altamente flexibles. Por ello, los diagramas de las figuras 2.13 y 2.14 son tan parecidos y muestran como las señales tienden a aparecer en lugares muy determinados de los espectros en ausencia de estructura. Sin embargo, un análisis en detalle pone en evidencia que los valores de δ_{rc} obtenidos a partir de los δ experimentales de la α -sinucleína difieren de los teóricos obtenidos a partir de pentapéptidos lineales. En todos los casos analizados, el conjunto de señales que provienen de los valores experimentales, aunque están corregidos por efecto de secuencia, es más disperso que el teórico. Así, por ejemplo, en el diagrama que representa las correlaciones $\text{C}'\text{-C}_\alpha$ (sin tener en cuenta los datos correspondientes a Pro), las señales ocupan un área máxima de 128.21 ppm^2 para el caso de los δ_{rc} que procede de los valores de la sinucleína y 96.24

ppm^2 para los datos puramente teóricos, lo que da una idea de la mayor dispersión en el primer caso.

Esto nos permite decir que el estado desestructurado de las IDPs no es igual al asumido como random coil. O lo que es lo mismo, que en estas proteínas, el esqueleto mantiene unas tendencias conformacionales que no están promediadas en su totalidad y que hay poblaciones preferidas que pueden evidenciarse a través de estos valores.

De cualquier manera, la congestión espectral de las señales en los experimentos de RMN de estos sistemas es muy grande y se hace preciso explotar la mínima posibilidad de aumentar la dispersión. Para ello, nada mejor que intentar sacar provecho de la característica que hemos demostrado de que se produce una mayor dispersión cuando se utilizan las correlaciones interresiduales, en lugar de correlaciones intrarresiduales. Por otra parte sabemos que las correlaciones que involucran los heteronúcleos del esqueleto C_α y el N_H son las más útiles para la asignación espectral (Takeuchi et al. 2008; Takeuchi et al. 2010). Por ello se puede pensar en diseñar versiones de experimentos 2D del tipo inter-CACO e inter-CAN

La información correspondiente al experimento inter-CACO puede obtenerse fácilmente adquiriendo el plano 2D CACO del experimento 3D CANCO (Bertini et al. 2004; Bermel et al. 2005b), basado en la detección del $^{13}\text{C}'$. El experimento 3D CANCO correlaciona los desplazamientos químicos de los núcleos C_α con los de los dos nitrógenos vecinos y por ende con los C' adyacentes a estos nitrógenos (Bertini et al. 2004a; Bermel et al. 2005b).

Pero más interesante parece el experimento inter-CAN, que proporciona (ver figura 2.21) una mejor dispersión de las señales con respecto a las otras posibilidades y que no había sido descrito hasta la fecha. El experimento inter-CAN puede ser empleado para estudiar y asignar los desplazamientos químicos tanto de proteínas bien plegadas como de proteínas desordenadas. Además, el hecho de observarse un carbono de tipo alifático hace que la secuencia sea prometedora para sistemas grandes que deban deuterarse ya que las señales no se verán afectadas por procesos de ensanchamiento ni en el caso de trabajar a campos magnéticos elevados.

4.2. Aplicación a la α -sinucleína

La α -sinucleína ha sido la proteína empleada en el presente trabajo como modelo de proteína intrínsecamente desordenada (Eliezer et al. 2001). Estas proteínas desestructuradas poseen una elevada flexibilidad conformacional lo que les permite adoptar diferentes estructuras y, por tanto, reconocer diferentes macromoléculas y/o ligandos conservando la especificidad en el reconocimiento de los mismos (Palsdottir & Hunte 2004; Ulmer et al. 2005). Esta gran versatilidad y sus grandes posibilidades de interacción las ha convertido en un centro de gran atención en la actualidad.

Si comparamos los dos experimentos de detección de ^{13}C con información interresidual, CON e inter-CAN, podemos comprobar su complementariedad y la ventaja de registrar este último en términos de asignación espectral. El experimento CON (Bermel et al. 2006) para la sinucleína conduce a espectros con los picos relativamente dispersos, incluso tratándose de una proteína desplegada, y las señales de los residuos de prolina están claramente identificadas debido al particular valor de desplazamiento químico del nitrógeno del esqueleto alrededor de 137 ppm, figura 2.22. Es evidente pues que este tipo de experimento resulta rico en información en comparación con aquella que obtenemos del espectro de correlación 2D ^1H - ^{15}N que se toma como referencia en la detección de ^1H .

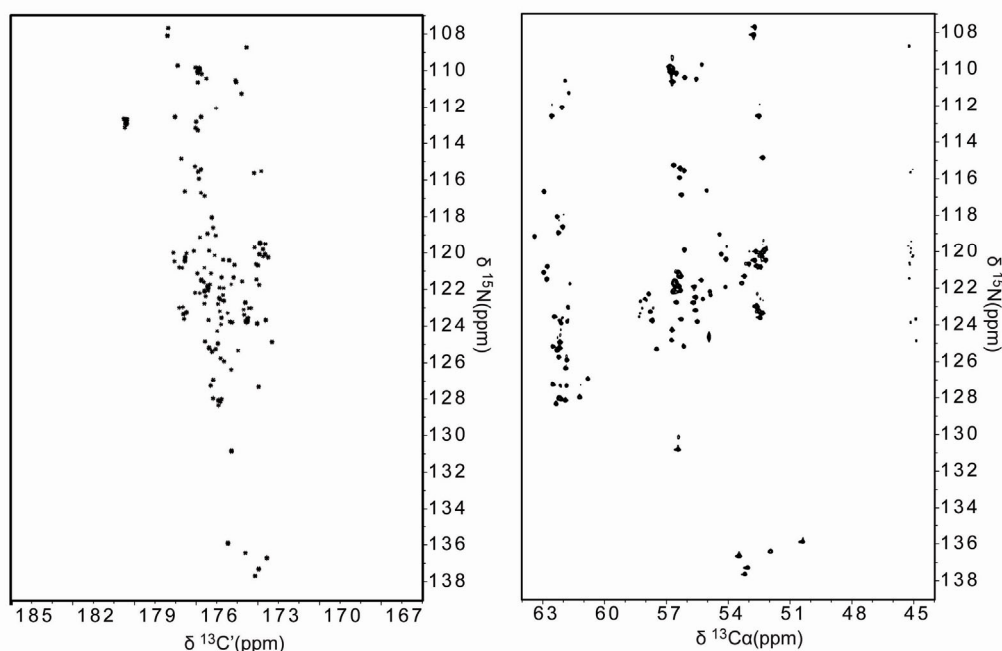


Figura 2.22. Experimentos CON e inter-CAN adquiridos en un espectrómetro Bruker Avance de 16.4 T sobre una muestra 1.0 mM de α -sinucleína. Ambos espectros poseen las mismas escalas en hercios tanto para el ^{13}C , 20 ppm, como para el ^{15}N , 32 ppm.

Pero es evidente que la dispersión en el espectro inter-CAN es mucho mejor que en el CON. La distribución general de las señales del espectro 2D inter-CAN experimental es muy parecida a la representación de las señales de correlación correspondientes obtenida a partir de los desplazamientos químicos δ_{rc} calculados en el apartado 3.1. La comparación puede verse en la figura 2.23.

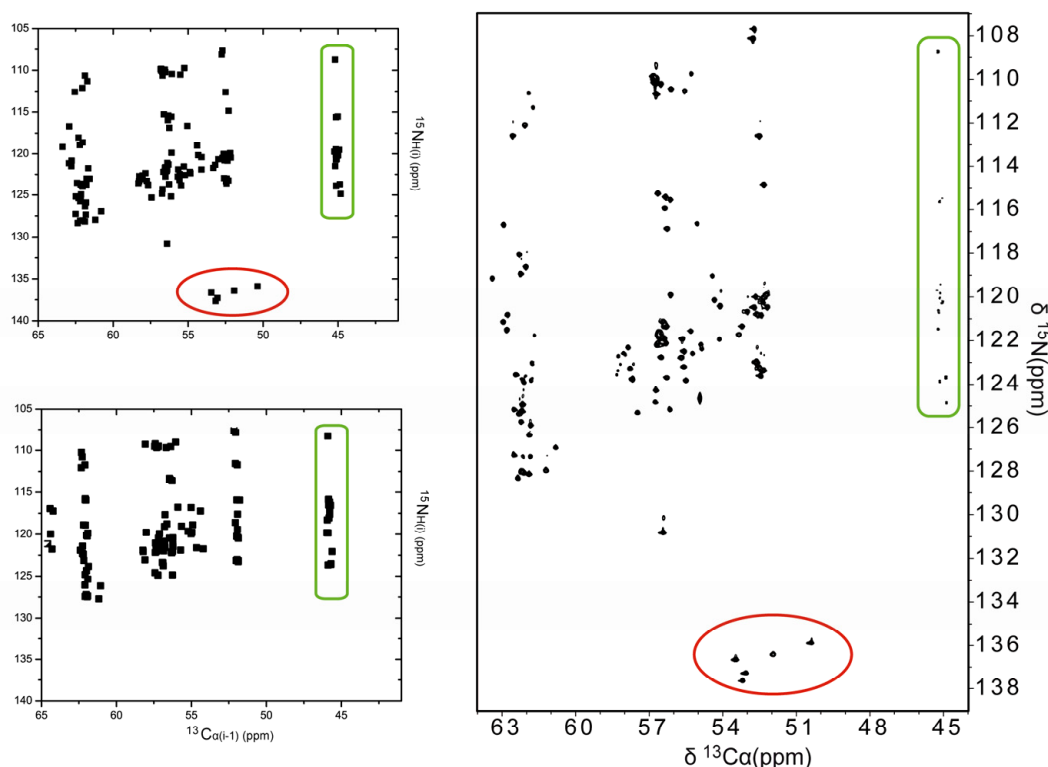


Figura 2.23. A la derecha experimento inter-CAN adquirido en un espectrómetro Bruker Avance de 16.4 T sobre una muestra 1.0 mM de α -sinucleína. A la izquierda representación de los picos de correlación correspondientes a partir de los valores de δ_{rc} empleando bien los que derivan de los valores experimentales (panel superior), bien los teóricos, en la panel inferior. En círculo rojo se representan las señales procedentes de los residuos de Pro y en rectángulo verde las de los residuos de Gly.

Como resulta apreciable en la figura 2.23, el espectro 2D inter-CAN de la α -sinucleína contiene las señales de correlación de cada uno de los residuos, incluidas las prolinas. Así, en la parte central inferior (círculo rojo) encontramos las señales correspondientes a los $\text{C}_\alpha\text{-N}$ de las 5 prolinas (Pro108, Pro117, Pro120, Pro128 y Pro138). También se identifican claramente en la parte derecha del espectro (rectángulo verde) las señales de correlación $\text{C}_\alpha\text{-N}_\text{H}$ de las glicinas, la cuales tienen valores de desplazamientos químicos para los C_α muy representativos, próximos a 45 ppm.

Para hacer una evaluación cuantitativa de la dispersión alcanzada, hemos usado la distancia media entre dos picos vecinos que es una medida bien aceptada para evaluar esta propiedad en un espacio de n-dimensiones (Clark and Evans, 1954). El valor obtenido es de 0.72 ppm en el espectro inter-CAN, mientras que este valor baja hasta solo 0.42 ppm en el caso de intra-CAN. Este dato justifica claramente el trabajo para el objetivo que se había planteado.

El empleo de NUS junto a la detección de ^{13}C nos ha permitido mejorar la calidad de los espectros registrados. Ambos enfoques pueden utilizarse tanto en proteínas plegadas como en IDPs, pero su empleo es particularmente interesante en estas últimas que muestran en general poca estabilidad conformacional.

5. CONCLUSIONES

En este capítulo se han analizado exhaustivamente las ventajas y desventajas de la detección directa en ^{13}C sobre la detección de ^1H . Esto es interesante para el diseño de nuevas secuencias de pulsos que permitan optimizar el estudio de sistemas flexibles a nivel de resolución atómica.

En primer lugar se ha comprobado como los desplazamientos químicos experimentales de las proteínas intrínsecamente desordenadas, aun cuando se corrijan por efectos de la secuencia en las posiciones ± 1 y ± 2 , no reflejan fielmente el estado random coil. Ello demuestra que las IDPs pueblan algunos nichos conformacionales de forma preferente. Así, los valores de δ_{rc} obtenidos a partir de los datos experimentales de IDPs no deben usarse para representar el estado puramente desplegado.

Se ha demostrado mediante la simulación de los espectros heteronucleares ^{13}C - ^{13}C : CACO y ^{13}C - ^{15}N : CON y CAN, que el solapamiento espectral se reduce cuando se ponen de manifiesto las correlaciones interresiduales frente a las intrarresiduales. A partir de esta observación, se deduce que una solución simple a la obtención de la información del tipo inter-CACO a partir el registro del plano correspondiente del experimento 3D-CANCO con detección de ^{13}C propuesto por Bermel y colaboradores, puede ser un buen complemento para la asignación espectral.

Además, se ha diseñado una nueva secuencia de pulsos para poder realizar el experimento 2D inter-CAN, basado en la detección directa de $^{13}\text{C}_\alpha$. Se ha propuesto como mejor solución para conseguir el desacoplamiento heteronuclear ^{13}C - ^{13}C de los carbonos alfa, la metodología IPAP en dos bloques que permite desacoplar $\text{C}_\alpha\text{-C}'$ y $\text{C}_\alpha\text{-C}_\beta$.

Esta secuencia se propone como una herramienta adicional, que puede ser utilizada junto a otros experimentos de detección de ^1H y ^{13}C , para la asignación de espectros de RMN de cualquier tipo de proteínas. Sin embargo, las ventajas que ofrece son mucho mayores en el caso de proteínas o regiones desestructuradas.

6. BIBLIOGRAFÍA

- Aitio H, Permi P (2000) Semi-constant-time HMQC (SCT-HMQC-HA) for the measurement of $(3)\text{J}(\text{H}(\text{N}))(\text{H}(\alpha))$ couplings in $(15)\text{N}$ -labeled proteins. *J Magn Reson* 143, 391-396.
- Andersson P, Weigelt J, Otting G (1998) Spin-state selection filters for the measurement of heteronuclear one-bond coupling constants. *Journal of biomolecular NMR* 12, 435-441.
- Arinaitwe K, Kiremire BT, Muir DC et al. (2012) Atmospheric concentrations of polycyclic aromatic hydrocarbons in the watershed of Lake Victoria, East Africa. *Environmental science & technology* 46, 11524-11531.
- Arnesano F, Banci L, Bertini I et al. (2001) Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. *Biochemistry* 40, 1528-1539.
- Banci L, Bertini I, Ciofi-Baffoni S et al. (2008) A structural-dynamical characterization of human Cox17. *The Journal of biological chemistry* 283, 7912-7920.
- Banci L, Bertini I, Del Conte R et al. (1999) The solution structure of a monomeric, reduced form of human copper,zinc superoxide dismutase bearing the same charge as the native protein. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry* 4, 795-803.
- Bermel W, Bertini I, Csizmek V et al. (2009) H-start for exclusively heteronuclear NMR spectroscopy: the case of intrinsically disordered proteins. *J Magn Reson* 198, 275-281.
- Bermel W, Bertini I, Duma L et al. (2005a) Complete assignment of heteronuclear protein resonances by protonless NMR spectroscopy. *Angew Chem Int Ed Engl* 44, 3089-3092.
- Bermel W, Bertini I, Felli IC et al. (2012) Speeding up sequence specific assignment of IDPs. *Journal of biomolecular NMR* 53, 293-301.
- Bermel W, Bertini I, Felli IC et al. (2006a) Novel ^{13}C direct detection experiments, including extension to the third dimension, to perform the complete assignment of proteins. *J Magn Reson* 178, 56-64.
- Bermel W, Bertini I, Felli IC et al. (2006b) Protonless NMR experiments for sequence-specific assignment of backbone nuclei in unfolded proteins. *Journal of the American Chemical Society* 128, 3918-3919.
- Bermel W, Bertini I, Felli IC et al. (2007) A method for C(alpha) direct-detection in protonless NMR. *J Magn Reson* 188, 301-310.
- Bermel W, Bertini I, Felli IC et al. (2005b) A selective experiment for the sequential protein backbone assignment from 3D heteronuclear spectra. *J Magn Reson* 172, 324-328.
- Bermel W, Bertini I, I.C. F et al. (2006c) ^{13}C -detected *protonless* NMR spectroscopy of proteins in solution. *Progress in nuclear magnetic resonance spectroscopy* 48, 25-45.
- Bermel W, Bruix M, Felli IC et al. (2013) Improving the chemical shift dispersion of multidimensional NMR spectra of intrinsically disordered proteins. *Journal of biomolecular NMR* 55, 231-237.
- Bermel W, I.C. F, R. K et al. (2008) ^{13}C direct-detection biomolecular NMR. *Concepts Magn Reson* 32A, 183-200.
- Bertini I, Duma L, I.C. F et al. (2004a) A heteronuclear direct detection NMR experiments for protein backbone assignment. *Angew Chem Int ed* 43, 2257-2259.
- Bertini I, Felli IC, Gonnelli L et al. (2011) High-resolution characterization of intrinsic disorder in proteins: expanding the suite of $(13)\text{C}$ -detected NMR spectroscopy experiments to determine key observables. *Chembiochem : a European journal of chemical biology* 12, 2347-2352.
- Bertini I, Felli IC, Kummerle R et al. (2004b) $(13)\text{C}$ - $(13)\text{C}$ NOESY: a constructive use of $(13)\text{C}$ - $(13)\text{C}$ spin-diffusion. *Journal of biomolecular NMR* 30, 245-251.

- Bertini I, Felli IC, Kummerle R et al. (2004c) ^{13}C - ^{13}C NOESY: an attractive alternative for studying large macromolecules. *Journal of the American Chemical Society* 126, 464-465.
- Bertini I, Lalli D, Mangani S et al. (2012) Structural insights into the ferroxidase site of ferritins from higher eukaryotes. *Journal of the American Chemical Society* 134, 6169-6176.
- Bertini I, Luchinat C, Parigi G (2001) *Solution NMR of Paramagnetic Molecules*. Elsevier.
- Bertini I, Luchinat C, Parigi G et al. (2005) NMR spectroscopy of paramagnetic metalloproteins. *Chembiochem : a European journal of chemical biology* 6, 1536-1549.
- Bodenhausen G, D.J. R (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 69, 185-189.
- Coggins BE, Zhou P (2006) Polar Fourier transforms of radially sampled NMR data. *J Magn Reson* 182, 84-95.
- Coggins BE, Zhou P (2007) Sampling of the NMR time domain along concentric rings. *J Magn Reson* 184, 207-221.
- Duma L, Hediger S, Lesage A et al. (2003) Spin-state selection in solid-state NMR. *J Magn Reson* 164, 187-195.
- Eletsky A, Moreira O, Kovacs H et al. (2003) A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using ^{13}C spectroscopy. *Journal of biomolecular NMR* 26, 167-179.
- Eliezer D, Kutluay E, Bussell R, Jr. et al. (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. *Journal of molecular biology* 307, 1061-1073.
- Englander SW, Downer NW, Teitelbaum H (1972) Hydrogen exchange. *Annual review of biochemistry* 41, 903-924.
- Eriksen JL, Dawson TM, Dickson DW et al. (2003) Caught in the act: alpha-synuclein is the culprit in Parkinson's disease. *Neuron* 40, 453-456.
- Felli IC, Brutscher B (2009) Recent advances in solution NMR: fast methods and heteronuclear direct detection. *Chemphyschem : a European journal of chemical physics and physical chemistry* 10, 1356-1368.
- Fyfe CA (1983) *Solid State NMR for Chemists*. C.F.C. Press.
- Goddard TD, Kneller DG (2005) *Sparky 3*, University of California, San Francisco.
- Grzesiek S, A. B (1992) An Efficient Experiment For Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins. *Journal of Magnetic Resonance* 99(1), 201-207.
- Harris RK (1986) *Nuclear Magnetic Resonance Spectroscopy*. Longman London.
- Hoch JC, Maciejewski MW, Filipovic B (2008) Randomization improves sparse sampling in multidimensional NMR. *J Magn Reson* 193, 317-320.
- Huang C, Ren G, Zhou H et al. (2005) A new method for purification of recombinant human alpha-synuclein in *Escherichia coli*. *Protein expression and purification* 42, 173-177.
- Ishii Y, Yesinowski JP, Tycko R (2001) Sensitivity enhancement in solid-state (^{13}C) NMR of synthetic polymers and biopolymers by (^1H) NMR detection with high-speed magic angle spinning. *Journal of the American Chemical Society* 123, 2921-2922.
- Keller RTJ (2004) *The computer aided resonance assignment tutorial*, Cantina Verlag.
- Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of biomolecular NMR* 49, 139-149.
- Kjaergaard M, Poulsen FM (2011) Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of biomolecular NMR* 50, 157-165.
- Kovacs H, Moskau D, Spraul M (2005a) Cryogenically cooled probes-a leap in NMR technology. *Progress NMR Spectroscopy* 46, 131-155.
- Kovacs H, Moskau D, Spraul M (2005b) Cryogenically cooled probes-a leap in NMR technology. *Progress in nuclear magnetic resonance spectroscopy* 46, 131-155.

- Kumar A, Rani Grace RC, Madhu PK (2000) Cross-correlations in NMR. *Progress in nuclear magnetic resonance spectroscopy* 37, 191-319.
- Lee D, Hilty C, Wider G et al. (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* 178, 72-76.
- LeMaster DM, Richards FM (1988) NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteration. *Biochemistry* 27, 142-150.
- Lindon JC, Holmes E, Nicholson JK (2005) *Metabonomics Techniques and Applications to Pharmaceutical Research & Development*. *Pharm Res* 45, 109.
- Maciejewski MW, Stern AS, King GF et al. (2006) Nonuniform Sampling in Biomolecular NMR. *Modern Magnetic Resonance* 2, 1305-1311.
- Marion D (2006) Processing of ND NMR spectra sampled in polar coordinates: a simple Fourier transform instead of a reconstruction. *Journal of biomolecular NMR* 36, 45-54.
- Marion D, Ikura M, R. T et al. (1989) Rapid Recording of 2D NMR Spectra without Phase Cycling. Application to the Study of Hydrogen Exchange in Proteins. *J Magn Reson* 85, 393-399.
- Markley JL, Bax A, Arata Y et al. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids--IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and nucleic acid structures determined by NMR spectroscopy. *European journal of biochemistry / FEBS* 256, 1-15.
- Markus MA, Dayie KT, Matsudaira P et al. (1994) Effect of deuteration on the amide proton relaxation rates in proteins. *Heteronuclear NMR experiments on villin 14T*. *Journal of magnetic resonance Series B* 105, 192-195.
- Merlino A, Mazzarella L, Carannante A et al. (2005) The importance of dynamic effects on the enzyme activity: X-ray structure and molecular dynamics of onconase mutants. *The Journal of biological chemistry* 280, 17953-17960.
- Miclet E, Williams Jr DC, Clore GM et al. (2004) Relaxation-optimized NMR spectroscopy of methylene groups in proteins and nucleic acids. *Journal of the American Chemical Society* 126, 10560-10570.
- Mittag T, Forman-Kay JD (2007) Atomic-level characterization of disordered protein ensembles. *Current opinion in structural biology* 17, 3-14.
- Mulder FA, Spronk CA, Slijper M et al. (1996) Improved HSQC experiments for the observation of exchange broadened signals. *Journal of biomolecular NMR* 8, 223-228.
- Murata T, Yamato I, Kakinuma Y et al. (2005) Structure of the rotor of the V-Type Na^+ -ATPase from *Enterococcus hirae*. *Science* 308, 654-659.
- Oschkinat H, Müller T, Dieckmann T (1994) Protein Structure Determination with Three- and Four- Dimensional NMR Spectroscopy. *Angew Chem Int ed* 33, 277-293.
- Ottiger M, Delaglio F, Bax A (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* 131, 373-378.
- Palsdottir H, Hunte C (2004) Lipids in membrane protein structures. *Biochimica et biophysica acta* 1666, 2-18.
- Pannetier N, Houben K, Blanchard L et al. (2007) Optimized 3D-NMR sampling for resonance assignment of partially unfolded proteins. *J Magn Reson* 186, 142-149.
- Pervushin K, Riek R, Wider G et al. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12366-12371.
- Pervushin KV, Wider G, Wuthrich K (1998) Single Transition-to-single Transition Polarization Transfer (ST2-PT) in $[\text{15N}, \text{1H}]$ -TROSY. *Journal of biomolecular NMR* 12, 345-348.
- Post CB (2003) Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Current opinion in structural biology* 13, 581-588.
- Robert D (1995) Carbon-13 Nuclear Magnetic Resonance Spectrometry. *Methods in Lignin Chemistry Book*, 250-273.

- Santoro J, King G (1992) A Constant-Time 2D Overbroaden Experiment for Inverse Correlation of Isotopically Enriched Species. *J Magn Reson* 97, 202-207.
- Sattler M, Fesik SW (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* 4, 1245-1249.
- Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in nuclear magnetic resonance spectroscopy* 34, 93-158.
- Selenko P, Wagner G (2007) Looking into live cells with in-cell NMR spectroscopy. *Journal of structural biology* 158, 244-253.
- Shaka AJ, Barker PB, Freeman R (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *Journal of Magnetic Resonance* 64, 547-552.
- Shaka AJ, Keeler J, Freeman R (1983) Evaluation of a new broadband decoupling sequence: WALTZ-16. *Journal of Magnetic Resonance* 53, 33-340.
- States DJ, Habekorn RA, Ruben DJ (1982) A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants. *J Magn Reson* 48.
- Takeuchi K, Frueh DP, Hyberts SG et al. (2010) High-resolution 3D CANCA NMR experiments for complete mainchain assignments using C(alpha) direct detection. *Journal of the American Chemical Society* 132, 2945-2951.
- Takeuchi K, Sun ZY, Wagner G (2008) Alternate ^{13}C - ^{12}C labeling for complete mainchain resonance assignments using C alpha direct-detection with applicability toward fast relaxing protein systems. *Journal of the American Chemical Society* 130, 17210-17211.
- Tamiola K, Acar B, Mulder FA (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *Journal of the American Chemical Society* 132, 18000-18003.
- Tompa P (2002) Intrinsically unstructured proteins. *Trends in biochemical sciences* 27, 527-533.
- Tugarinov V, Hwang PM, Ollerenshaw JE et al. (2003) Cross-correlated relaxation enhanced ^1H - ^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *Journal of the American Chemical Society* 125, 10420-10428.
- Ulmer TS, Bax A, Cole NB et al. (2005) Structure and dynamics of micelle-bound human alpha-synuclein. *The Journal of biological chemistry* 280, 9595-9603.
- Wishart DS, Bigam CG, Yao J et al. (1995) ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *Journal of biomolecular NMR* 6, 135-140.
- Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. J Wiley&Sons, Inc New York.

Capítulo III

Péptidos boomerang

Capítulo III:

Péptidos boomerang

1. INTRODUCCIÓN

1.1. Los glicosaminoglicanos

Los glicosaminoglicanos, GAG, son polisacáridos de cadena larga y no ramificada, cuyos pesos moleculares están en el intervalo de 10 a 100 kDa. Están compuestos por la repetición de disacáridos, formados por un azúcar ácido, que suele ser el ácido D-glucurónico (GlcA) o bien ácido L-idurónico (IdoA), y un amino azúcar, que puede ser D-glucosamina (GlcN) o D-galactosamina (GalN), figura 3.1. Estos azúcares presentan grados variables de sulfatación, generalmente en las posiciones 3 y 6 del anillo de piranosa. El alto contenido de grupos ácidos presentes en los GAG hace que estén cargados negativamente. Su importancia radica en que están involucrados en multitud de procesos biológicos que incluyen, entre otros, cicatrización de heridas (Peplow 2005), anticoagulación (Casu et al. 2004), señalización celular, desarrollo y angiogénesis (Kovensky 2009), progresión tumoral y metástasis (Liu et al. 2002); e incluso pueden jugar un papel importante en enfermedades relacionadas con depósitos amiloides (Kisilevsky et al. 2007).

Dentro de la familia de los GAG destaca la heparina, un polisacárido sulfatado formado por la repetición de 12 a 50 veces del disacárido ácido L-idurónico-N-acetil-D-glucosa, que se emplea en medicina como anticoagulante para el tratamiento de trombosis, tromboflebitis y embolias (Casu et al. 2004). La heparina presenta efectos secundarios indeseados, como complicaciones sanguíneas o trombocitopenia (disminución de la cantidad de plaquetas que circulan en el torrente sanguíneo). Este hecho unido a los avances en el conocimiento del

proceso en cascada de la coagulación, y a la separación de los componentes de distintos pesos moleculares de la heparina, han conducido hacia el desarrollo de análogos de heparina con menor peso molecular. Estos compuestos, que tienen propiedades químicas y biológicas mejor definidas (Choay et al. 1981), han ido desplazando a la heparina como anticoagulante en los tratamientos clínicos. Uno de los análogos de última generación de la heparina es el Fondaparinux (Gandhi & Mancera 2010), comercializado por la compañía GlaxoSmithKline con el nombre de Arixtra®. El compuesto Fondaparinux, figura 3.1, es un pentasacárido con numerosos grupos cargados negativamente que se emplea como agente antitrombótico y actúa como inhibidor selectivo del factor indirecto Xa mediante su unión selectiva a “serine protease inhibitor antithrombin III” (Turpie 2004; Turpie et al. 2004).

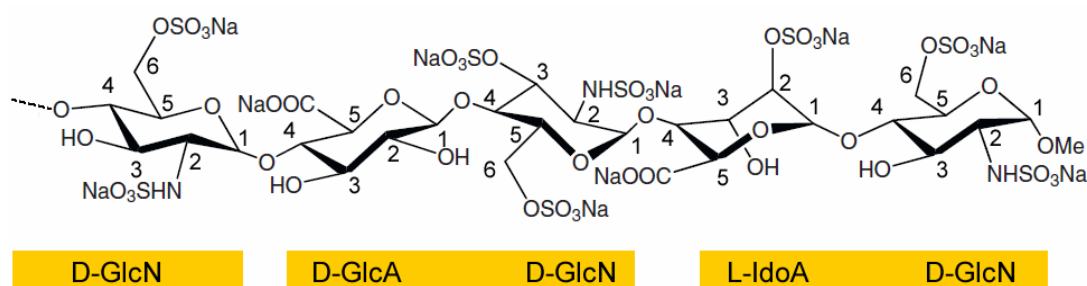


Figura 3.1. Estructura química del pentasacárido Fondaparinux SR9010107, Org31540 (Arixtra®). Se indica la numeración de los carbonos.

1.2. Reconocimiento proteína/heparina

Se han descrito numerosas proteínas que unen heparina. Estas proteínas presentan una gran diversidad de funciones, tales como factores de crecimiento, trombinas, proteínas virales, lo que las hace de interés desde el punto de vista farmacológico. A pesar de ello, la información disponible acerca de los requerimientos moleculares responsables del reconocimiento específico proteína/GAG y proteína/heparina es aún escasa. No obstante, a partir del análisis de las estructuras y secuencias de proteínas que unen heparina y otros GAG, y de modelizaciones moleculares se han propuesto y descrito varios motivos de secuencia y estructurales para el reconocimiento de GAG.

Los primeros motivos de reconocimiento de GAG descritos fueron las secuencias XBBXBX y XBBBXXBX, donde B es un residuo básico (Lys o Arg) y X un residuo neutro o hidrófobo, definidos a partir del análisis de cuatro proteínas que unen GAG (Cardin & Weintraub 1989). Estudios de modelización mostraron que en ciertas conformaciones los residuos básicos de estas secuencias se disponen en una orientación que facilita su interacción con los grupos sulfato ($-\text{OSO}_3$) cargados negativamente de la heparina. Así, si la secuencia XBBXBX se encuentra en una conformación de hebra β , los aminoácidos básicos

se alinean hacia una de las caras de la hebra, mientras que los aminoácidos hidrófobos o neutros se disponen hacia la otra, donde se encontraría el núcleo de la proteína. Del mismo modo, si la secuencia de XBBBXXBX adopta estructura de hélice α , los aminoácidos básicos se posicionan a uno de los lados de la hélice, mientras que el resto apuntan hacia el otro lado, hacia el interior de la proteína. Utilizando este modelo, Sobel y sus colaboradores propusieron una nueva secuencia consenso de unión entre una proteína y la heparina, XBBBXXBBBXXBBX (Sobel et al. 1992).

Estos motivos de secuencia se emplearon como sondas para predecir la capacidad de unión a heparina de una proteína a partir de su secuencia. Sin embargo, algunas de las proteínas que se unen a heparina, descubiertas posteriormente, no contienen estos motivos de secuencia. Dado que residuos básicos alejados en la secuencia pueden estar próximos en la estructura terciaria, el factor determinante de la capacidad de unión a heparina es la orientación espacial de los residuos básicos con independencia de su proximidad en la secuencia. Mediante el examen de proteínas con sitios de unión a heparina contiguos linealmente, Margalit y colaboradores encontraron que una distancia de aproximadamente 20 Å entre los aminoácidos básicos es importante para la interacción, independientemente de que la estructura del sitio de unión sea una hélice α o una hebra β (Margalit et al. 1993). Como los residuos básicos se disponen en lados opuestos del elemento de estructura secundaria, los autores proponen que la heparina se enrolla alrededor de estos residuos cargados positivamente, por lo que su unión puede inducir cambios conformacionales en las proteínas.

Los sitios de unión a heparina suelen encontrarse en la superficie externa de las proteínas y corresponden a bolsillos poco profundos de carga positiva. Por lo tanto, la topología del sitio de unión y no sólo la secuencia, es importante para la unión a heparina. Mediante el análisis estructural de los sitios de unión a heparina en varios factores de crecimiento (FGF-1, FGF-2 y TGF β -1) se ha identificado el motivo TXXBXXTBXXXTBB, donde T es un giro, B un residuo básico de arginina o lisina (o en ocasiones una glutamina) y X un residuo polar o hidrófobo (Hileman et al. 1998).

La importancia de los aminoácidos comunes a los dominios de unión a heparina de diversas proteínas, así como de la separación de los aminoácidos básicos en la unión a heparina se ha examinado mediante el estudio de péptidos modelo (Capila & Linhardt 2002). Utilizando bibliotecas de heptapéptidos se ha encontrado que los péptidos enriquecidos en Arg y Lys, pero no His, se unen con mayor afinidad a heparina y al sulfato de heparán (Capila & Linhardt 2002). Los péptidos con alta afinidad por heparina y sulfato de heparán también

están enriquecidos en otros aminoácidos polares como la Ser (Caldwell et al. 1996). Los dominios de unión a heparina conocidos también contienen aminoácidos capaces de formar enlaces de hidrógeno (Gln y Asn). Además, entre los aminoácidos básicos, la Arg se une 2.5 veces más fuertemente a heparina que la Lys. Esto puede ser consecuencia del fuerte enlace de hidrógeno entre el grupo guanidinio de la arginina y el grupo sulfato de la heparina (Fromm et al. 1995). El efecto del patrón y el espaciamiento de los aminoácidos básicos en los sitios de unión a heparina y sulfatos de heparán se ha estudiado también en dos series de péptidos sintéticos, $-RRG_mRR-$ y $-RRRG_mR-$ (donde R corresponde a Arg, G a Gly y m al número de glicinas que va de 0 a 5) (Fromm et al. 1997). Se encontró que una interacción fuerte con heparina requiere un sitio de unión con alta densidad de carga positiva, mientras que en el caso de los sulfatos de heparán, que tienen un menor contenido de grupos sulfato, la interacción es más fuerte cuando los residuos básicos están más espaciados.

Más recientemente (Torrent et al. 2012), mediante el análisis de las estructuras de numerosas proteínas que unen heparina se ha propuesto un nuevo motivo estructural de unión a heparina, denominado clip CPC' (catiónico-polar-catiónico). Este motivo está formado por un residuo polar (Asn, Gln, Tyr, Ser o Thr) y dos residuos catiónicos (Lys o Arg) con la disposición espacial mostrada en la figura 3.2, que viene definida por las siguientes distancias: 6.0 ± 1.8 Å (PC), 11.6 ± 1.6 Å (PC') y 11.4 ± 2.4 Å (CC') para los C_α y 6.0 ± 1.9 Å (PC), 10.6 ± 1.8 Å (PC') y 10.7 ± 2.0 Å (CC') para el centro de gravedad de las cadenas laterales.

Los residuos de Arg y Lys proporcionan los primeros contactos a través de enlaces de hidrógeno mientras que los residuos polares, por su parte, son los que pueden determinar el reconocimiento preciso de GAGs. Por lo tanto, estos resultados sugieren que en la unión de proteínas con heparina se conserva más un patrón estructural que un patrón en la secuencia y el motivo CPC' podría considerarse el requisito estructural mínimo para tal efecto.

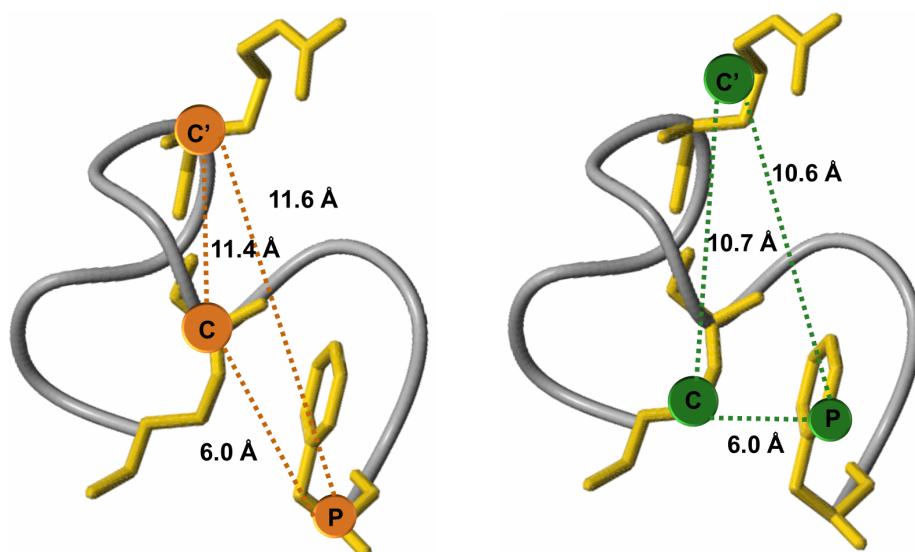


Figura 3.2. Representación del motivo CPC' formado por los residuos Lys5-Tyr1-Arg8 del péptido YW12 código pdb 2O0S. El esqueleto peptídico está representado en gris y las cadenas laterales de los residuos del motivo CPC' en amarillo. En la imagen de la izquierda las esferas naranjas representan los C_{α} de los residuos que forman el motivo CPC' mientras que en la imagen de la derecha, las esferas verdes indican el centro de gravedad de las cadenas laterales de los residuos Lys5-Tyr1-Arg8.

1.3. Lipopolisacáridos

Los lipopolisacáridos, LPS, también llamados endotoxinas, son el principal constituyente de la membrana externa de las bacterias Gram negativas (Raetz & Whitfield 2002). Su estructura y disposición de cargas en las distintas unidades que lo forman proporcionan un medio propicio para el plegamiento de las proteínas de membrana.

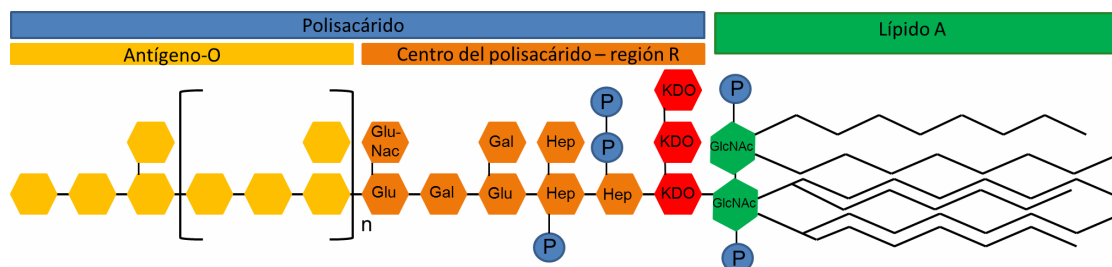


Figura 3.3. Esquema de la estructura de los LPS. El lípido A está formado por un disacárido compuesto por dos unidades de N-acetilglucosamina (GlcNAc) fosforiladas unidas mediante un enlace éster a las cadenas de ácidos grasos. La región R o centro del polisacárido está constituida por seis monosacáridos de hexosa. Y por último, el antígeno-O, que es la parte de los LPS que queda hacia el medio extracelular, actúa como receptor para los bacteriófagos.

Los LPS son polímeros complejos que se componen de dos partes, el lípido A y una cadena de polisacáridos, figura 3.3. El lípido A está formado por dos N-acetilglucosaminas (GlcNAc) fosforiladas, las cuales están unidas a seis cadenas saturadas de ácido graso. A

partir de estas cadenas los LPS se unen a la membrana externa de las bacterias Gram-negativas. En la cadena de polisacáridos se distinguen dos zonas, el centro del polisacárido o región R, un oligosacárido que está unido a una de las GlcNAc del lípido A, y el antígeno-O, una serie de 3 a 5 monosacáridos que se repiten hasta 40 veces en algunos casos y que es diferente en función a las especies.

1.4. Los péptidos antimicrobianos “ β -boomerang” adoptan el motivo de unión a heparina clip CPC’ en micelas de lipopolisacáridos (LPS)

En el año 2007 el grupo del profesor Bhattacharjya (Bhattacharjya et al. 2007) diseñó un dodecapéptido de secuencia YVLWKRKRMIFI, que fue nombrado como YW12, que presenta actividad antimicrobiana, figura 3.2. Este péptido carece de estructura ordenada en disolución acuosa, pero de acuerdo con los experimentos RMN de tipo tr-NOESY (donde se observan señales de transferencia de NOE) adopta una estructura definida de β -boomerang, figura 3.4, al interaccionar con micelas de lipopolisacáridos. Curiosamente, esta estructura de β -boomerang contiene el motivo clip CPC’ de reconocimiento de heparina. Hay que tener en cuenta que los lipopolisacáridos presentan características químicas comunes con los GAG (ver apartados 1.2 y 1.3).

El diseño del péptido YW12 se basó en la estructura cristalina de la proteína FhuA, que se encuentra en la membrana externa de la bacteria *Escherichia coli*, en su complejo con LPS. Esta proteína tiene estructura de barril β (código de acceso pdb 1QFG) (Ferguson et al. 1998). Para el diseño de este péptido se consideraron los residuos de la proteína FhuA que tienen contactos con los LPS, figura 3.5, cuyas distancias interatómicas proteína/LPS son iguales o inferiores a 4.0 Å, y que se encuentran distribuidos a lo largo de segmentos discontinuos de lámina β en la cadena peptídica (Ferguson et al. 1998). Considerando que en la estructura cristalina del complejo FhuA/LPS los grupos bifosfato del lípido A están involucrados en múltiples enlaces de hidrógeno y puentes salinos con los residuos cargados positivamente Lys439, Lys441, Arg384 y Lys357 de FhuA, la secuencia de YW12 incluye cuatro residuos básicos en la parte central del péptido, Lys5-Arg6-Lys7-Arg8, que deben formar el motivo de unión a los LPS. Aparte de las interacciones iónicas y polares, la estructura cristalina FhuA/LPS presenta contactos de van der Waals (distancias interatómicas de 4.0 Å) entre dos grupos de residuos aromáticos y/o alifáticos (Tyr284, Leu300, Val282 y Phe235; y Phe302, Phe355 y Phe380) y las cadenas alquílicas de los LPS, figura 3.5. El péptido YW12 intenta reproducir el empaquetamiento de estos grupos hidrófobos mediante su respectiva incorporación en los extremos N- y C-terminales de su secuencia. Así, los primeros cuatro residuos del péptido YW12, Tyr1-Val2-Leu3-Trp4, corresponden a los

residuos Tyr284, Val282, Leu300 y Phe235 de FhuA. El residuo en posición 4 es Trp en lugar de Phe para tener una sonda de fluorescencia en el péptido. Se seleccionó esa posición para introducir el Trp para que se encuentre lo más próximo posible a los residuos con carga positiva. Los tres últimos residuos del péptido YW12, Ile10-Phe11-Ile12, corresponden a los residuos aromáticos Phe302, Phe355 y Phe380 de FhuA, donde dos de las Phe se han sustituido por Ile, por la mayor propensión a la conformación de lámina β que presentan los residuos ramificados de Ile en un entorno de membrana (Bishop et al. 2001). En posición 9 del péptido YW12 se colocó una Met con el fin de facilitar el análisis sin ambigüedades de los espectros de RMN. Además, los residuos de Met se encuentran con frecuencia en péptidos activos de la membrana antimicrobiana, indicando su posible implicación en la interacción con los lípidos (Zasloff 2002).

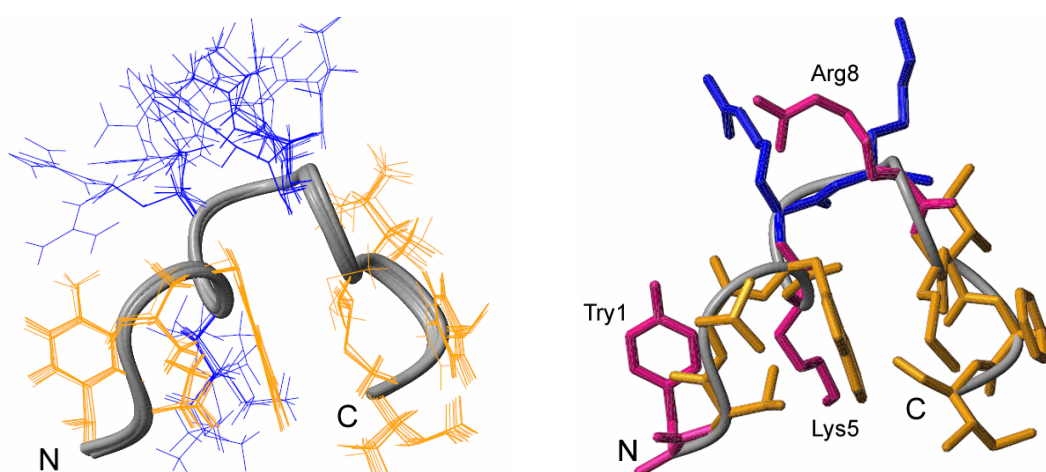


Figura 3.4. Estructuras del péptido YW12 en presencia de LPS, código pdb 2O0S, representadas con el programa MOLMOL (Koradi et al. 1996). A la izquierda, familia de estructuras más favorables desde el punto de vista energético, con el esqueleto peptídico representado en gris, las cadenas de los residuos con carga positiva en azul y el resto de cadenas laterales en naranja. A la derecha, una de las estructuras representada siguiendo el mismo código de colores, salvo en las cadenas laterales de los residuos que forman el motivo CPC; Lys5-Tyr1-Arg8, que se resaltan en rosa.

Estudios de fluorescencia del triptófano demuestran que el péptido YW12 muestra interacciones selectivas sólo con las micelas formadas por lípidos cargados negativamente (incluyendo SDS, POPG y LPS) y no se detectan interacciones significativas en micelas zwitterionicas como las de DPC. Mediante CD en el UV lejano se observa la presencia de giros β o lámina β en los péptidos con micelas con carga negativa mientras que no aparecen transiciones estructurales en micelas de DPC (Bhattacharjya et al. 2007).

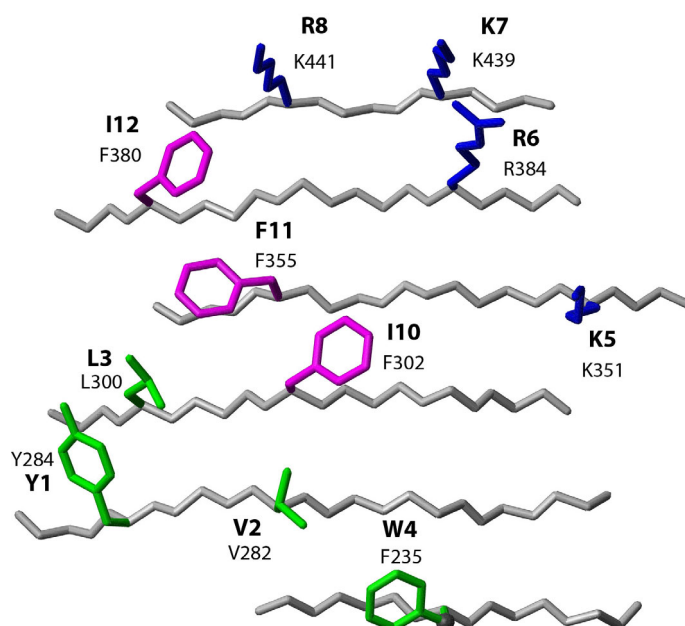


Figura 3.5. Residuos en la superficie de interacción de la proteína FhuA de *Escherichia coli* con LPS, código PDB 1QFG (Ferguson et al. 2000). Los números sin resaltar indican las posiciones de los residuos que tienen contacto con los LPS en la estructura FhuA/LPS y los números en negrita indican la posición a la que se corresponden en los péptidos estudiados.

La estructura del péptido YW12 en el complejo con LPS (Bhattacharjya et al. 2007) que se ha determinado usando la espectroscopía Tr-NOE (Clare & Gronenborn 1982; Post 2003), no presenta ninguna estructura secundaria regular, figura 3.4: la mitad N-terminal, Val2-Arg6, tiene una conformación extendida o de bucle, mientras que en la C-terminal, Lys7-Ile12, se distinguen dos giros β . La orientación de las cadenas laterales de los residuos hidrófobos y aromáticos se encuentra, en general, bien definida, mientras que las cadenas laterales de los cuatro residuos con carga positiva, Lys5, Arg6, Lys7 y Arg8, están menos definidas. La mayoría de los residuos hidrófobos y aromáticos están orientados hacia un lado mientras que los residuos centrales cargados se posicionan hacia el lado contrario, dando lugar a una estructura anfipática. Las cadenas laterales de los residuos Val2, Leu3 y Trp4 del extremo N-terminal se empaquetan con las cadenas laterales de los residuos Met9, Phe11 e Ile12 del extremo C-terminal de YW12 formando el “cluster” hidrófobo que une los LPS como se muestra la figura 3.6. Las interacciones entre las cadenas laterales del Trp4 y la Met9 posiblemente juegan un papel importante conectando los grupos hidrófobos de los extremos N- y C-terminales. Las cadenas laterales de los residuos Tyr1 e Ile10 se sitúan en la parte superior de la superficie hidrofóbica.

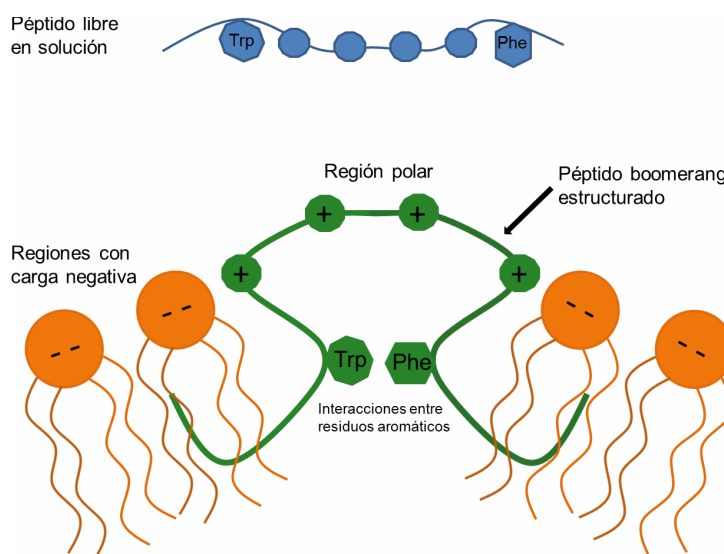


Figura 3.6. Modelo propuesto del modo de interacción y plegamiento del péptido YI12WF y sus derivados con LPS (Bhunia et al. 2009). En la parte superior esquema del péptido desestructurado en solución. El centro de los residuos con carga positiva puede interaccionar con las regiones polares y los grupos fosfatos de la cabeza de los LPS. Los residuos aromáticos, críticos en el empaquetamiento propuesto en la estructura unida a LPS, proporcionan la orientación de los brazos hidrófobos en los extremos N- y C-terminales del péptido en la conformación activa.

A partir del diseño del péptido YW12 (Bhattacharjya et al. 2007), Bhunia y colaboradores diseñaron nuevos péptidos mediante la sustitución de los residuos Trp4 y Met9 por otros residuos hidrófobos (Bhunia et al. 2009). Todos los péptidos derivados de YW12 muestran una estructura anfipática en los complejos con LPS, donde los residuos con carga positiva se encuentran en el bucle definiendo la superficie catiónica de la molécula, mientras que los residuos hidrófobos de los extremos N- y C- terminales adoptan conformaciones extendidas o hebras β formando la superficie hidrofóbica, figura 3.6. La estructura tipo β que se da por el empaquetamiento de los residuos aromáticos 4 y 9 se asemeja a un “boomerang” o estructura tipo β -boomerang. Con anterioridad se había descrito una estructura de hélice boomerang (Han et al. 2001) en un dominio de fusión de 20 residuos, donde el empaquetamiento de los residuos localizados en las posiciones (i) e (i+5) de un bucle hidrófilo definen la orientación entre el segmento helicoidal precedente y el que le sigue.

Entre los péptidos diseñados (Bhunia et al. 2009) destaca el péptido YI12WF, cuya secuencia es igual a YW12, pero con Phe en la posición 9 en lugar de Met. Este cambio de la Met por el aminoácido aromático, mejora el empaquetamiento con el Trp4 en los complejos péptido/LPS (Bhattacharjya et al. 2007). Por ello, se cree que esta sustitución podría ser la responsable de la definición de la estructura, dando lugar a la gran superficie hidrofóbica del péptido cuando se encuentra en presencia de LPS con la consiguiente mejora en la neutralización de las actividades endotóxicas y antimicrobianas. .

1.5. Objetivo

Las interacciones proteína/GAG y proteína/heparina juegan un papel fundamental en numerosos procesos biológicos. En particular, la heparina se emplea en medicina clínica como agente anticoagulante. Recientemente (Torrent et al. 2012), se ha descrito el motivo estructural de clip (o grapa) CPC' (catiónico-polar-catiónico), conservado en numerosas proteínas con capacidad de unión a heparina, como el principal contribuyente en la formación de complejos proteína/heparina. Este motivo de grapa CPC' se encuentra también presente en los péptidos antimicrobianos de tipo β -boomerang (Bhunja et al. 2009), por lo que sería de esperar que fueran capaces de unirse a heparina y otros GAG, y constituirían un modelo sencillo para el análisis de los factores responsables de la interacción proteína/heparina. Con esta hipótesis en mente, el objetivo de este capítulo es determinar si los péptidos antimicrobianos tipo β -boomerang efectivamente unen heparina, y en caso afirmativo, confirmar si el motivo clip CPC' es el responsable de dicha unión. Para ello, se examinará la importancia para la interacción de todos los residuos catiónicos, así como del residuo polar y de su posición en la secuencia haciendo uso de la técnica de RMN en disolución.

Dado que la heparina es una mezcla de polisacáridos de diverso peso molecular, y que su interacción con el péptido YI12WF (en este trabajo denominado YI12), cabeza de la serie, condujo a la precipitación de la muestra (apartado 2.3), se ha empleado un análogo de heparina de menor tamaño y de estructura química bien definida, el pentasacárido Fondaparinux, figura 3.1. Además, este análogo se emplea actualmente en clínica para evitar algunos de los efectos secundarios no deseados de la heparina.

2. MATERIALES Y MÉTODOS

2.1. Síntesis de los péptidos

Los péptidos fueron sintetizados en el grupo del Profesor David Andreu de la Universitat Pompeu Fabra de Barcelona mediante la tecnología de fase sólida y fueron purificados por HPLC de fase inversa.

2.2. Requisitos de la muestra

Dado que las características de la muestra influyen en gran medida en la calidad y en el tipo de los espectros que pueden adquirirse, hay que seleccionar las condiciones óptimas para realizar los experimentos. La finalidad es encontrar las condiciones que mejor reproduzcan aquellas en las que la proteína o el péptido sean biológicamente activos y en las que a su vez sea posible la adquisición de espectros de RMN de buena calidad. Para ello hay que tomar en consideración los siguientes aspectos, que se tendrán en cuenta en éste y en los sucesivos capítulos de este trabajo:

- La concentración de la proteína o de los péptidos. Para obtener una buena relación señal/ruido, las muestras han de tener una concentración del orden de 1 mM. Este requerimiento se debe al hecho de que la RMN es intrínsecamente una técnica de baja sensibilidad (Jiménez et al. 2003).
- La estabilidad de la muestra. La proteína o péptido objeto de estudio debería permanecer inalterado durante el tiempo de adquisición de los espectros, que va desde los minutos (monodimensionales), a las horas (experimentos bidimensionales), e incluso a los días (experimentos 3D y 4D). Hay que elegir las condiciones de pH, la fuerza iónica, el tampón, la temperatura, etc. en las que la proteína o el péptido sean más estables, y que más se aproximen a las condiciones biológicas en que realizan su función.
- Dado que hay que trabajar con concentraciones “altas” será necesario que la proteína o el péptido sean solubles a la concentración requerida. La adición de cosolventes, como el 2,2,2-trifluoroetanol, TFE, puede solventar los problemas de insolubilidad.

2.3. Preparación de las muestras

Todas las muestras preparadas contienen una concentración aproximada de 0.1 mM DSS (2,2-dimetil-2-silapentano-5-sulfonato sódico) como referencia interna de los desplazamientos químicos de ^1H .

2.3.1. Péptidos en solución acuosa

Para la preparación de las muestras de cada uno de los péptidos en agua se pesaron aproximadamente 1 mg de péptido liofilizado y se disolvieron en 0.5 mL de una solución acuosa $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v que contenía tampón HEPES 2 mM. El pH se ajustó a aproximadamente 5.0 mediante la adición de alícuotas de DCl o NaOD y se midió con un microelectrodo de vidrio Hamilton de 3 mm de diámetro.

2.3.2. Titulaciones con heparina

Se preparó una disolución de heparina (Sigma-Aldrich código H6279) a concentración 18 mM. Se añadieron alícuotas de 5, 10, 25 y 50 μL de esta disolución de heparina a una muestra del péptido Y112 en disolución acuosa a una concentración 1 mM, y se adquirieron espectros de RMN tras cada adición. Se examinó el efecto del pH en el intervalo 2.1-12.7. Las relaciones molares péptido/heparina empleadas fueron desde 1.0/0.0 hasta 1/1.80.

2.3.3. Titulaciones con Fondaparinux

Cada vial inyectable Arixtra® (de GlaxoSmithKline) contiene 2.5 mg del pentasacárido Fondaparinux, figura 3.1 (Gandhi & Mancera 2010), en 0.5 mL de solución acuosa, lo que corresponde a una concentración de 2.6 mM. La adición de 50 μL de esta disolución a una muestra de 1 mg del péptido Y112 en 500 μL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v a pH 5.0 condujo a la precipitación total del péptido. El precipitado se redisolvió mediante la adición de 200 μL del cosolvente 2,2,2-trifluoroetanol deuterado, TFE (Cambridge Isotope Laboratories, DLM-2318-1). La proporción de TFE en la muestra fue de 28.6 % en volumen. En todas las muestras el pH se ajustó a 5.0 mediante la adición de alícuotas de DCl o NaOD según correspondiese.

Las titulaciones de los demás péptidos con Fondaparinux se realizaron siguiendo este mismo protocolo, y todos los espectros de RMN se adquirieron en presencia del cosolvente.

2.3.4. Péptidos en presencia del cosolvente TFE

Puesto que la presencia del TFE afecta a los desplazamientos químicos del péptido, para poder analizar la interacción de los péptidos con Fondaparinux se adquirieron espectros de cada uno de los péptidos aislados en presencia de la misma proporción de TFE (28.6 % en volumen) y al mismo pH (5.0). Para esto se disolvió 1 mg de péptido en 500 μL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v, y se añadieron 200 μL de TFE.

2.4. **Adquisición de los espectros de RMN**

La adquisición de los espectros de RMN se realizó en un espectrómetro de RMN Bruker Avance 600, con un campo magnético de 14.1 Teslas que opera a 600.13 MHz para el protón, equipado con una criosonda TXI-“Criosonda Triple Inversa”-(^1H , ^{13}C , ^{15}N).

Para todas las muestras se adquirieron los experimentos bidimensionales homonucleares, ^1H - ^1H COSY (COrelated SpectroscopY) (Aue et al. 1976), ^1H - ^1H TOCSY (Total Correlated SpectroscopY) (Braunschweiler & Ernst 1983) y ^1H - ^1H NOESY (Nuclear Overhauser Effect SpectroscopY) (Jeener et al. 1979; Kumar et al. 1980; Macura & Ernst 1980), utilizando secuencias estándar de pulsos, eliminándose la señal del agua mediante presaturación. El tiempo de mezcla empleado para los experimentos ^1H - ^1H NOESY fue de 150 ms y para los espectros ^1H - ^1H TOCSY de 60 ms.

En el caso de los péptidos en disolución acuosa se adquirieron los espectros bidimensionales ^1H - ^1H ROESY (Rotating frame nuclear Overhauser Effect SpectroscopY) (Bax & Davis 1985).

Los experimentos se llevaron a cabo a 25 °C y la temperatura de la muestra dentro de la sonda de RMN se calibró utilizando una muestra de metanol (Duerst & Verbach 1965).

Para comprobar la estabilidad de los péptidos durante el transcurso de los experimentos se adquirieron espectros monodimensionales de ^1H al principio y al final de las diferentes series de espectros bidimensionales.

2.5. **Procesado de los espectros de RMN**

Los espectros mono y bidimensionales se procesaron con el programa TopSpin 3.0 (Bruker Biospin, Karlsruhe, Alemania) , a los cuales se les aplicó la transformada de Fourier y se realizó la corrección de línea base en la dimensión directa e indirecta según fuera necesario en cada caso.

2.6. Asignación espectral

Para llevar a cabo la asignación espectral completa se utilizó el procedimiento estándar desarrollado por Wüthrich (Wüthrich 1986) y se empleó el programa SPARKY (Goddard 2005). En primer lugar, a partir del análisis de los experimentos ^1H - ^1H COSY y ^1H - ^1H TOCSY, se identifican todos los protones pertenecientes al mismo sistema de espín y se clasifican como pertenecientes a un determinado tipo de aminoácido. En el experimento ^1H - ^1H COSY a partir del protón $^1\text{H}_\text{N}$ se observa la correlación con el protón $^1\text{H}_\alpha$, de este con $^1\text{H}_\beta$ y así sucesivamente, hasta completar el sistema de espín de que se trate. En el experimento ^1H - ^1H TOCSY se observan picos de cruce entre todos los protones pertenecientes a un determinado sistema de espín. Cada sistema de espín se asocia a un determinado tipo de aminoácido, si dicho aminoácido no se repite en la secuencia del péptido, que es conocida, queda ya asignado. Sin embargo, en este nivel, no es posible distinguir entre aminoácidos iguales que se repitan en la secuencia, ni entre residuos con el mismo sistema de espín y valores de desplazamiento químico característicos similares, tales como la pareja Asp y Asn o los aminoácidos aromáticos Phe, Tyr y Trp, todos ellos de tipo AMX .

Para poder asignar a que residuo corresponde cada uno de los sistemas de espín identificados, se utiliza el espectro ^1H - ^1H NOESY en el cual se observan señales de correlación NOE entre protones cercanos en el espacio (a distancia inferior a aproximadamente 5 Å). Los espectros ^1H - ^1H NOESY muestran tanto señales intrarresiduales como interresiduales, pudiendo estas últimas ser secuenciales y no secuenciales. Los NOEs secuenciales, es decir, los picos de cruce entre el protón $^1\text{H}_\alpha$ del residuo (i) y el protón H_N de un residuo (i+1), entre protones amídicos consecutivos, y entre el protón H_N y los protones de la cadena lateral del residuo anterior, son los que permiten la identificación de los aminoácidos de manera consecutiva según el orden de la secuencia, lo que se denomina asignación secuencial. Los NOEs intrarresiduales corroboran la asignación realizada con los espectros ^1H - ^1H TOCSY y ^1H - ^1H COSY, e incluso amplían la información, como en el caso de los aminoácidos aromáticos para los que se observa una señal de correlación entre los protones $^1\text{H}_\beta$ y los $^1\text{H}_\delta$ del anillo aromático. Las señales no secuenciales corresponden a los picos de cruce entre protones cercanos en el espacio pero pertenecientes a aminoácidos no consecutivos, por lo que son fundamentales para el cálculo de la estructura tridimensional.

2.7. Determinación de la estructura secundaria

Los elementos de estructura secundaria de péptidos y proteínas se pueden identificar de forma cualitativa utilizando los desplazamientos conformacionales de los protones $^1\text{H}_\alpha$, de los carbonos $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$, y mediante los NOEs secuenciales y de medio alcance.

Los desplazamientos conformacionales se definen como la diferencia entre el valor del desplazamiento químico observado experimentalmente y su valor en ausencia de estructura o random coil; ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm). En este trabajo se han tomado como valores de referencia de random coil los medidos en hexapéptidos sin estructura (Wishart et al. 1995a; Wishart et al. 1995b). Las hélices α se caracterizan por tener desplazamientos conformacionales negativos para los protones $^1\text{H}_\alpha$ y los carbonos $^{13}\text{C}_\beta$, y positivos para los $^{13}\text{C}_\alpha$, mientras que los desplazamientos conformacionales característicos de las hebras β son positivos para los protones $^1\text{H}_\alpha$ y los carbonos $^{13}\text{C}_\beta$, y negativos para los carbonos $^{13}\text{C}_\alpha$.

La estructura secundaria también se puede determinar a partir de los NOEs, ya que las correlaciones NOE características de los distintos tipos de estructuras secundarias, hélices α , láminas β y giros β , son diferentes. En la tabla 3.1 se muestran los valores para varios tipos de conformaciones.

Tipo de correlación	Señal	Intensidad (distancia protón-protón, Å)			
		Hélice α	Hélice 3_{10}	Lámina β antiparalela	Lámina β paralela
Correlaciones secuenciales	$d_{\alpha\text{N}(i,i+1)}$	media (3.5)	media (3.4)	fuerte (2.2)	fuerte (2.2)
	$d_{\text{NN}(i,i+1)}$	fuerte (2.8)	fuerte (2.6)	débil (4.3)	débil (4.2)
Correlaciones de medio alcance	$d_{\text{NN}(i,i+2)}$	débil (4.2)	débil (4.1)		
	$d_{\alpha\text{N}(i,i+2)}$	débil (4.4)	media (3.8)		
	$d_{\alpha\text{N}(i,i+3)}$	media (3.4)	media (3.8)		
	$d_{\alpha\text{N}(i,i+4)}$	débil (4.2)	---		
	$d_{\alpha\beta(i,i+3)}$	(2.5-4.4)	(3.1-5.1)		
	$d_{\alpha\text{N}(i,j)}$			media (3.2)	media (3.0)
	$d_{\text{NN}(i,j)}$			media (3.3)	media (4.0)
	$d_{\alpha\alpha(i,j)}$			fuerte (2.3)	débil (4.8)

Tabla 3.1. Correlaciones NOE características de hélices α , hélice 3_{10} , lámina β antiparalela y lámina β paralela.

2.8. Análisis de los resultados

La determinación de la estructura tridimensional de los péptidos de este trabajo no ha sido posible por la falta de NOEs no secuenciales en los espectros ^1H - ^1H NOESY adquiridos. Esto ha impedido obtener las restricciones de distancia necesarias para el cálculo de estructura. Por ello, la interacción de los péptidos con Fondaparinux se ha analizado mediante el cartografiado de desplazamientos químicos.

Dado que los desplazamientos químicos son muy sensibles al entorno químico de un núcleo, cambios en los mismos pueden indicar, no sólo cambios conformacionales en la molécula estudiada, sino también interacciones no covalentes con otras moléculas presentes en la muestra. Por ello, se pueden emplear los cambios de los desplazamientos químicos de una proteína originados por la adición de un ligando para localizar el sitio de unión, identificar las superficies de interacción, determinar la afinidad del ligando, y modelar la posible estructura del complejo proteína/ligando. Este método se conoce como perturbación o cartografiado de desplazamientos químicos (CSP). En proteínas, donde se suele disponer de la proteína etiquetada con ^{15}N o doblemente etiquetada con ^{15}N y ^{13}C , el cartografiado de desplazamientos químicos se realiza mediante la adquisición de espectros bidimensionales del tipo ^1H - ^{15}N HSQC y ^1H - ^{13}C HSQC, que permiten la asignación de los desplazamientos químicos de ^{15}N y ^{13}C , además de los de ^1H . Uno de los factores determinantes de la calidad de los espectros durante las titulaciones proteína/ligando, es la velocidad de intercambio entre la forma libre y unida, que en ocasiones origina un ensanchamiento de las señales de RMN.

En este caso, como los péptidos que se han empleado no estaban enriquecidos ni en ^{13}C ni en ^{15}N y además se han encontrado grandes problemas de solubilidad con las muestras, la interacción de los péptidos objeto de estudio de este capítulo, tabla 3.2, con Fondaparinux se ha analizado mediante el cartografiado de desplazamientos químicos de los átomos de ^1H , en concreto los correspondientes a los $^1\text{H}_\alpha$ y a los $^1\text{H}_\text{N}$.

Las diferencias observadas en los espectros con TFE y con TFE y Fondaparinux son pequeñas, y para llevar a cabo el cartografiado de las variaciones se ha empleado la fórmula:

$$\Delta\delta^{Av} = \sqrt{(\Delta\delta H_\alpha)^2 + (\Delta\delta N_H)^2}$$

$$\text{Donde: } \Delta\delta_{H_\alpha} = \delta H_\alpha (\text{péptido}+\text{TFE}+\text{Fondaparinux}) - \delta H_\alpha (\text{péptido}+\text{TFE})$$

$$\Delta\delta_{N_H} = \delta H_N (\text{péptido}+\text{TFE}+\text{Fondaparinux}) - \delta H_N (\text{péptido}+\text{TFE})$$

3. RESULTADOS

3.1. Diseño de los péptidos

Se ha tomado como modelo el péptido boomerang YI12WF diseñado por Bhunia y colaboradores (sección 1.4), y que aquí se denomina YI12, tabla 3.2. Con el fin de analizar las interacciones con heparina y su análogo Fondaparinux, se han diseñado siete péptidos. La nomenclatura empleada en los péptidos es el resultado de considerar el primer y el último aminoácido, seguido del número de residuos que contiene cada péptido. Los péptidos que difieren del cabeza de serie YI12 en un residuo no terminal, y que por tanto tienen extremos idénticos a YI12, se denominan como YI12 seguido del aminoácido cambiado y su posición. Así, M9 en el péptido YI12M9, que corresponde al péptido YW12 de Bhunia y colaboradores, indica que en la posición 9 hay una Met en lugar de una Phe, y de forma análoga los péptidos YI12Q5, YI12Q6, YI12Q7 y YI12Q8 tienen una Gln en las posiciones 5, 6, 7 y 8, respectivamente.

YI12	1Y	2V	3L	4W	5K	6R	7K	8R	9F	10I	11F	12I
YI12Q5	1Y	2V	3L	4W	<u>5Q</u>	6R	7K	8R	9F	10I	11F	12I
YI12Q6	1Y	2V	3L	4W	5K	<u>6Q</u>	7K	8R	9F	10I	11F	12I
YI12Q7	1Y	2V	3L	4W	5K	6R	<u>7Q</u>	8R	9F	10I	11F	12I
YI12Q8	1Y	2V	3L	4W	5K	6R	7K	<u>8Q</u>	9F	10I	11F	12I
FI12	<u>1F</u>	2V	3L	4W	5K	6R	7K	8R	9F	10I	11F	12I
FY12	<u>1F</u>	2V	3L	4W	5K	6R	7K	8R	9F	10I	11F	<u>12Y</u>
YY12	1Y	2V	3L	4W	5K	6R	7K	8R	9F	10I	11F	<u>12Y</u>
YI12M9	1Y	2V	3L	4W	5K	6R	7K	8R	<u>9M</u>	10I	11F	12I

Tabla 3.2. Secuencias de los péptidos estudiados. Las columnas marcadas muestran los residuos que pertenecen al motivo putativo CPC' de las posiciones 1, 5 y 8. Los residuos subrayados indican los residuos que se varían con respecto a YI12.

Para comprobar la importancia y función de los residuos catiónicos del centro de la secuencia se diseñaron los péptidos YI12Q5, YI12Q6, YI12Q7 e YI12Q8 en los que se sustituyen los residuos de Lys o Arg de las posiciones 5, 6, 7 y 8 por una Gln, respectivamente. El péptido FI12 se ha diseñado para determinar la importancia del residuo polar y componente del motivo CPC' de Tyr en la posición 1 y los efectos que se pueden producir al cambiarlo por un residuo no polar como la Phe. En el péptido FY12, como en el anterior, se elimina el residuo polar de la posición 1 pero se traslada a la posición 12 para estudiar cómo afecta el cambio de polaridad en el extremo C-terminal. El péptido YY12, con dos Tyr en las posiciones 1 y 12, debe permitir determinar el efecto en la estructura y capacidad de unión en presencia de un residuo polar adicional. Asimismo, se ha analizado el péptido YI12M9, originalmente diseñado por Bhunia et al. (sección 1.4).

3.2. Estructura secundaria de los péptidos en disolución acuosa y en presencia de TFE

La asignación secuencial específica para todos los aminoácidos de los péptidos se realizó combinando los espectros 2D ^1H - ^1H NOESY y ^1H - ^1H TOCSY, tal y como se describe en el apartado 2.6.

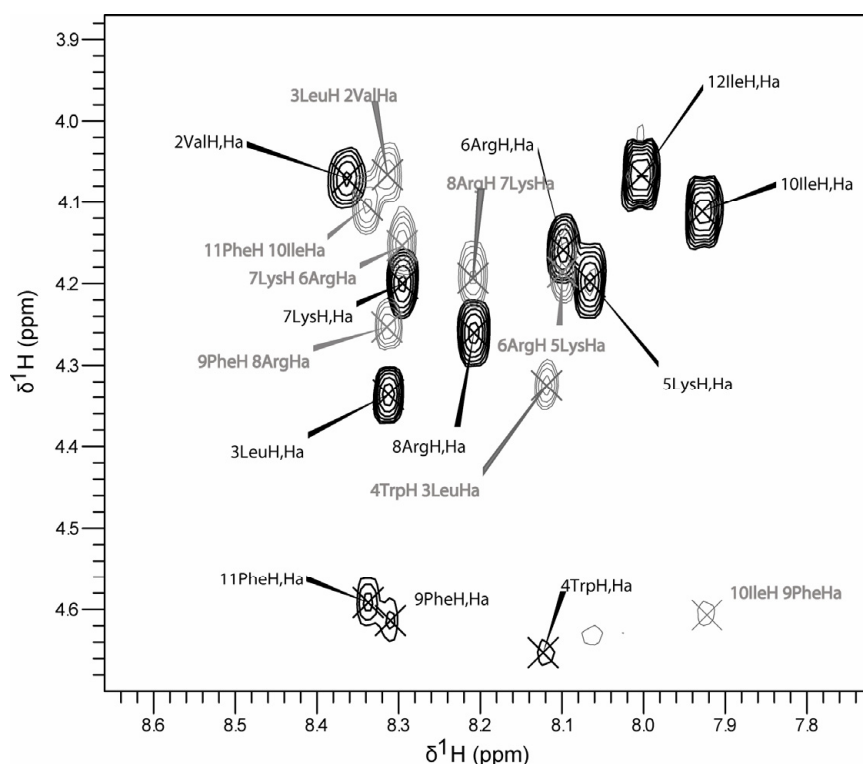


Figura 3.7. Sección de los experimentos bidimensionales ^1H - ^1H TOCSY (en negro) y ^1H - ^1H NOESY (en gris) del péptido YI12 en solución acuosa $\text{H}_2\text{O}/\text{D}_2\text{O}$ 9:1 v/v.

Aunque el análisis estructural por RMN del péptido YI12 en disolución acuosa ya había sido publicado (Bhunja et al. 2009) y las condiciones son las empleadas por Bhunja, se ha estudiado de nuevo para asegurarnos de que no había ninguna diferencia en las condiciones experimentales utilizadas en el péptido cabeza de serie YI12 y en los péptidos diseñados aquí, lo que facilita la comparación de su comportamiento en disolución. Los espectros ^1H - ^1H NOESY, del péptido YI12, figura 3.7, así como de todos los péptidos diseñados a partir de su secuencia, tabla 3.2, se caracterizan por la presencia de NOEs secuenciales e intrarresiduales débiles, y la ausencia de NOEs no secuenciales. Sólo en algún caso llega a observarse algún NOE poco intenso de tipo (i), (i+2). Para confirmar que la ausencia de NOEs no se debía a tiempos de correlación desfavorables de los péptidos, en los que la intensidad de los NOEs fuera muy pequeña, se realizaron experimentos ^1H - ^1H ROESY para

cada uno de los péptidos. En estos espectros tampoco se encontraron señales ROE adicionales de medio y largo alcance. Esta ausencia de NOEs no secuenciales indica, que en solución acuosa, las estructuras de todos los péptidos son altamente dinámicas. Este resultado está en concordancia con los resultados del grupo de Bhattacharjya (Bhattacharjya et al. 2007), que indicaban que el péptido Y112 carecía de estructura ordenada en disolución acuosa (Bhunia et al. 2009).

Los patrones de NOEs observados para todos los péptidos en presencia de TFE son similares a los encontrados en disolución acuosa, como puede verse para el péptido Y112 en la figura 3.8.

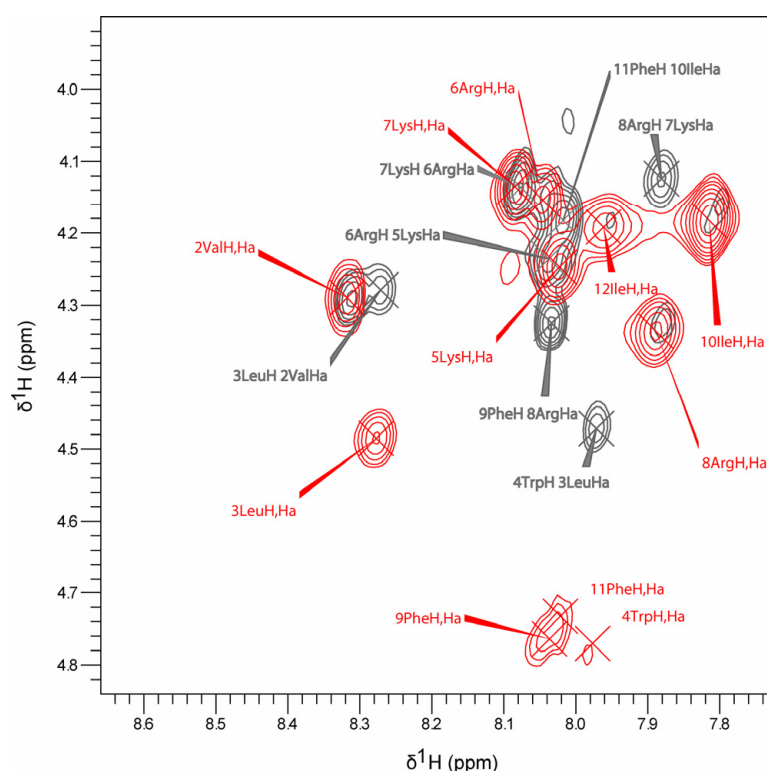


Figura 3.8. Péptido Y112 en solución con 28.6 % en volumen de TFE. Las señales rojas corresponden al experimento ^1H - ^1H TOCSY y las grises con el experimento ^1H - ^1H NOESY.

Dada la imposibilidad de comparar las posibles diferencias conformacionales de los péptidos a partir de los NOEs tanto en disolución acuosa como en presencia de TFE, se ha procedido a analizar las tendencias estructurales mediante los desplazamientos conformacionales ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm) (ver apartado 2.7). Los desplazamientos conformacionales de los protones H_α se han representado frente a la secuencia en las figuras 3.9, 3.10 y 3.11. En el perfil del péptido Y112 en disolución acuosa se observa que los residuos centrales, Lys5, Arg6, Lys7 y Arg8, presentan valores de $\Delta\delta$ negativos y más

grandes en valor absoluto que el resto ($|\Delta\delta| > 0.09$ ppm), que son indicativos de la formación de hélice o giro. En presencia de TFE, mientras que los residuos del segmento N-terminal (2-4) y los del extremo C-terminal (9-11) presentan valores de $\Delta\delta$ positivos y grandes, los residuos 5-8 mantienen los valores de $\Delta\delta$ negativos, figura 3.9. Este perfil sugiere que el péptido YI12 en presencia de TFE adopta una estructura de tipo horquilla β , hebra/giro/hebra. Esta estructura concuerda con la estructura de β -boomerang que puede considerarse como una horquilla β abierta en la que las hebras N- y C-terminales divergen, en lugar de ser prácticamente paralelas, como ocurre en las horquillas β canónicas (Sibanda et al. 1989). Estos resultados coinciden con los descritos por Bhunia que indican que el péptido adopta esta estructura en presencia de micelas de LPS (Bhunia et al. 2009).

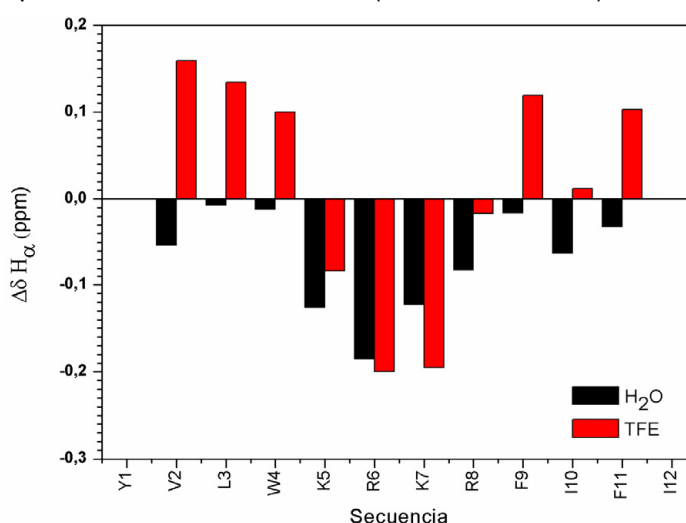


Figura 3.9. Perfiles de $\Delta\delta_{H\alpha}$ del péptido YI12 en disolución acuosa en negro ($\Delta\delta_{H\alpha} = \delta_{H\alpha}(\text{péptido}+\text{H}_2\text{O}) - \delta_{H\alpha}(\text{random coil})$) y en TFE en rojo ($\Delta\delta_{H\alpha} = \delta_{H\alpha}(\text{péptido}+\text{TFE}) - \delta_{H\alpha}(\text{random coil})$).

Los cuatro péptidos en los que uno de los residuos catiónicos centrales ha sido sustituido por el residuo no cargado Gln, tabla 3.2, presentan perfiles de desplazamientos conformationales de los protones H_α similares a los observados para el péptido YI12, figura 3.10, y por tanto, un comportamiento estructural análogo. Hay que mencionar que en el caso del péptido YI12Q7, los valores de H_α de los residuos Trp4 y Phe9 no pudieron ser asignados en presencia de TFE. Las diferencias más destacables se observan para los péptidos YI12Q6 e YI12Q8 en presencia de TFE, donde hay una disminución apreciable de la magnitud de los valores $\Delta\delta$ positivos de los residuos 2, 3, 9 y 11. Estos resultados indican que la estructura tipo horquilla β o bien no se forma o bien se forma en menor medida en estos péptidos que en el péptido cabeza de serie YI12, y que en los péptidos YI12Q5 y YI12Q7. Es interesante observar que los péptidos que conservan el comportamiento del cabeza de serie son los que presentan las sustituciones en los dos residuos de Lys originales en las posiciones 5 y 7, tabla 3.2.

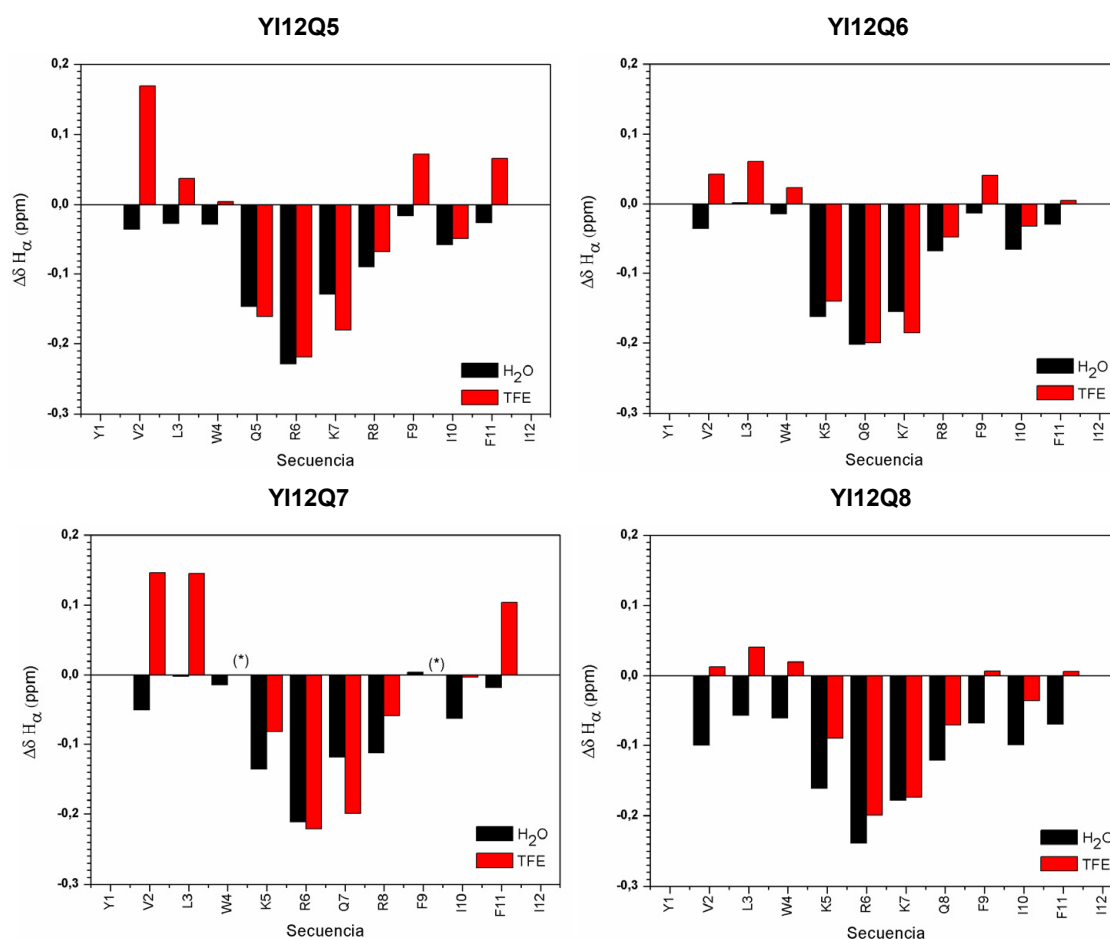


Figura 3.10. Diagrama de barras de los perfiles de $\Delta\delta_{H_\alpha}$ en disolución acuosa en negro y en TFE en rojo. Los residuos marcados con un (*) en YI12Q7 indican que esos valores no se han podido determinar.

Los perfiles de los desplazamientos conformacionales de los protones H_α de los péptidos FI12, FY12 e YY12, que tienen mutaciones en los residuos N- y C-terminales, tabla 3.2, se muestran en la figura 3.11. Los $\Delta\delta$ de estos péptidos tienen la misma tendencia que YI12 tanto en disolución acuosa como en presencia de TFE, aunque con diferencias en el caso de los péptidos YY12 y FY12, en los que el residuo 12 ha sido cambiado de Ile a Tyr. En estos dos péptidos es difícil hacer una evaluación de la formación de la hebra C-terminal basada en la comparación con el random coil debido a la presencia de dos residuos aromáticos consecutivos. Es bien conocido que en estos casos las corrientes de anillo pueden modificar sustancialmente los valores de δ de protones cercanos (Gomes & Mallion 2001) y modificar los perfiles de los desplazamientos conformacionales. En ambos casos, sin embargo, parece que en TFE debe de promoverse la estructura de β -boomerang ya que los residuos N-terminales 2-5 y el 9 presentan valores positivos de $\Delta\delta$.

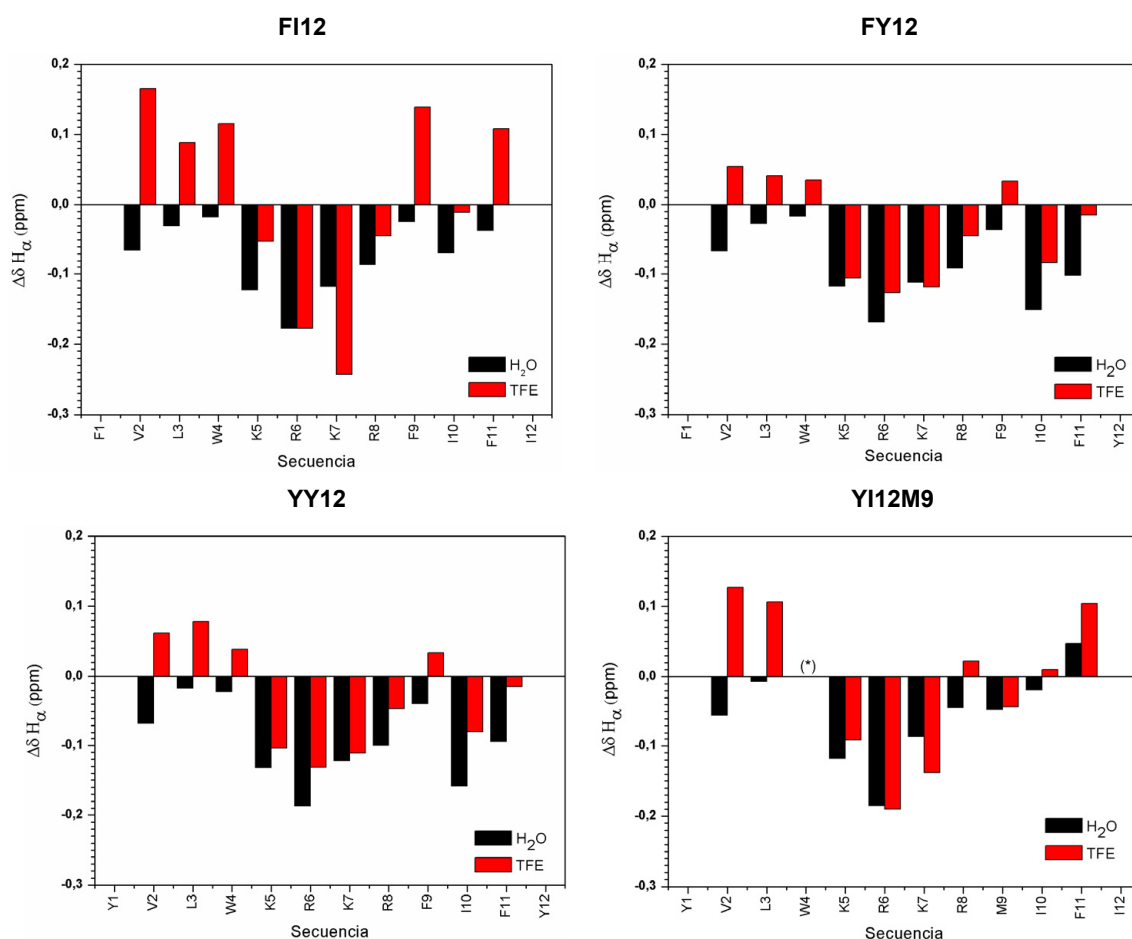


Figura 3.11. Representación de los desplazamientos químicos conformacionales de los péptidos FI12, FY12, YY12 e YI12M9 en negro en H₂O y en rojo en TFE. El residuo marcado con un (*) indica que ese valor no se ha podido determinar.

Los perfiles de desplazamientos conformacionales del péptido YI12M9, el diseñado originalmente por Bhattacharjya y colaboradores (Bhattacharjya et al. 2007) (ver apartado 1.4), siguen una tendencia similar a la observada para los demás péptidos, figura 3.11. Sin embargo, es el único péptido en el que el residuo Arg8 presenta un valor positivo en presencia de TFE, y el residuo 9 un valor negativo, lo que sugiere diferencias en la región del giro. El segmento C-terminal muestra diferencias con respecto a los demás péptidos, aunque el valor positivo y grande de $\Delta\delta$ de Phe11, unido a los valores positivos, pero muy pequeños en magnitud, de Arg8 e Ile10, sugiere que conserva la tendencia a formar conformaciones extendidas.

3.3. Titulación de YI12 con heparina

Para confirmar experimentalmente que el péptido YI12, que contiene el motivo CPC', es capaz de unir heparina, a una muestra del péptido en disolución acuosa se le fueron adicionando alícuotas de una disolución concentrada de heparina (ver apartado 2.3.2). Tras cada adición de heparina se adquirió un espectro de RMN monodimensional. Además de verificar la interacción entre el péptido YI12 y la heparina, el objetivo de esta titulación era encontrar las condiciones adecuadas a emplear para el estudio de las interacciones péptido/heparina en el resto de los péptidos diseñados, tabla 3.2. Tras la primera adición de 5 μ L de heparina, correspondiente a una proporción péptido/heparina de 1/0.18, en el espectro 1D sólo se veían las señales del péptido, figura 3.12. La intensidad de las señales correspondientes al péptido fue disminuyendo al ir aumentando la proporción de heparina. A partir de una relación molar péptido/heparina 1/1.20, comienzan a aparecer las señales de la heparina, pero desaparecen las correspondientes al péptido. Cuando se alcanzó la proporción 1/1.80 sólo se observaban las señales de la heparina. Estos resultados pueden explicarse si el complejo péptido/heparina es insoluble en el medio acuoso y precipita o forma agregados de alto peso molecular que no son visibles por RMN. Por estos motivos, la intensidad de las señales del péptido en el espectro de RMN va disminuyendo a medida que éste se une a la heparina, hasta desaparecer totalmente. A partir de este momento, empiezan a aparecer las señales del exceso de heparina no unida al péptido. Por tanto, se confirma la capacidad de unión del péptido YI12 a heparina, pero el hecho de que el complejo formado precipite o sea demasiado grande, impide obtener información acerca de la estructura del péptido en el complejo con heparina.

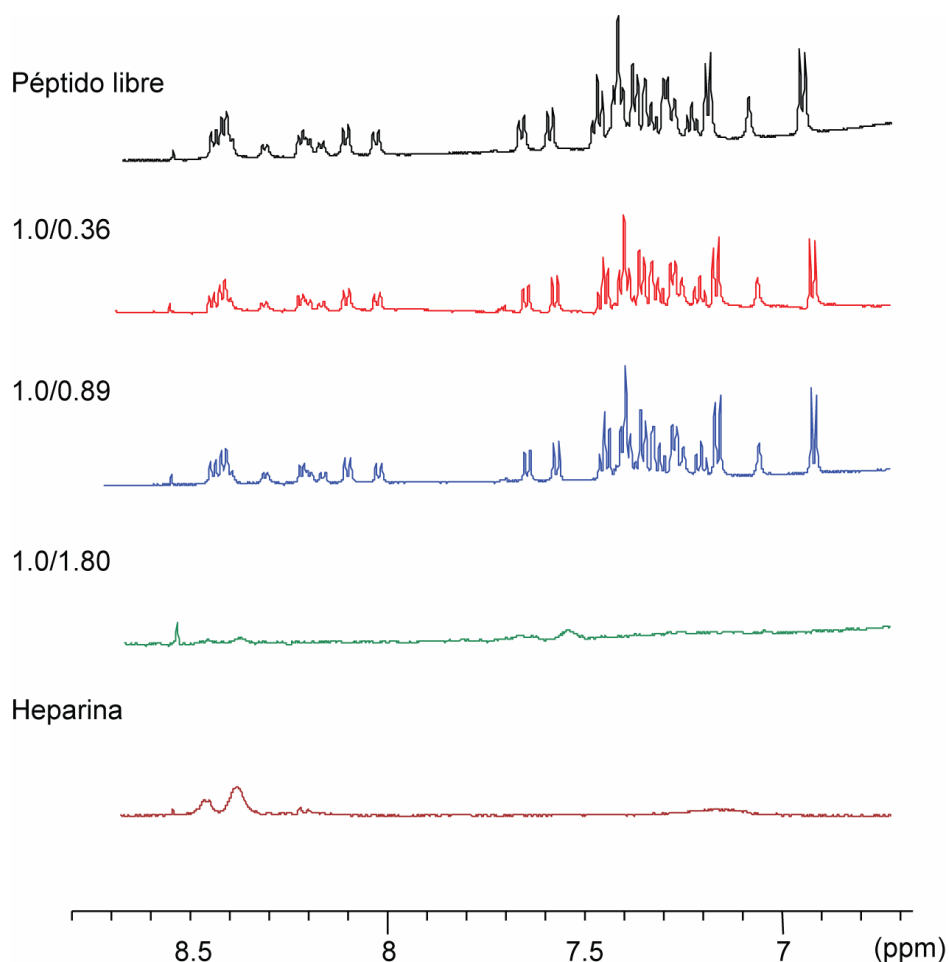


Figura 3.12. Región, de 6.8 a 8.7 ppm, de los espectros monodimensionales de ^1H del péptido Y112 a 25 °C y pH 5 a distintas relaciones molares péptido/heparina. De la parte superior a la inferior: en negro, el péptido libre; en rojo, con 10 μL de heparina; en azul, más 25 μL de heparina; en verde, más 50 μL de heparina y en granate, heparina sola.

3.4. Interacciones del péptido YI12 con Fondaparinux

Dado que la heparina comercial es químicamente heterogénea, y que el péptido YI12 al interactuar con heparina precipita, se pensó que una alternativa más prometedora podría ser la utilización de un análogo de heparina con una estructura química homogénea. Entre ellos, se consideró especialmente interesante el Fondaparinux, figura 3.1, que se emplea en medicina clínica (apartado 1.1). Desafortunadamente, al igual que en el caso de heparina, la adición de Fondaparinux a una muestra del péptido YI12 en una relación péptido/Fondaparinux 1/0.5 condujo a la precipitación del péptido. De nuevo, estos resultados indican que hay interacción entre el péptido y el análogo de heparina, pero la precipitación impide su estudio por RMN a nivel de resolución atómica. Este precipitado se redisolvió completamente mediante la adición de TFE, haciendo posible la adquisición de espectros de RMN bidimensionales. Aun así, la calidad de los espectros, en particular, los de tipo ^1H - ^1H NOESY no es buena, debido principalmente al ensanchamiento de las señales. Además, con el tiempo se vuelve a producir la precipitación. A pesar de los inconvenientes, la comparación de los parámetros de RMN obtenidos para el péptido YI12 en TFE en presencia y en ausencia de Fondaparinux puede proporcionar información acerca de la interacción YI12/Fondaparinux.

El perfil de los $\Delta\delta$ de los protones H_α que presenta el péptido YI12 en presencia de Fondaparinux y TFE, es muy similar al observado para el péptido por sí solo en TFE, figura 3.13. El hecho de que los valores de los desplazamientos químicos varíen poco indica que la conformación global del péptido no se altera por la presencia de Fondaparinux. Sin embargo, cuando se analizan los datos con detalle, se observan cambios pequeños en los desplazamientos químicos de algunas cadenas laterales, entre los que cabe destacar los correspondientes al anillo indólico de Trp4 y los de los anillos de Tyr1 y Phe9, como puede verse en la figura 3.14. Los cambios en los protones del esqueleto, H_N amida y H_α debidos a la presencia de Fondaparinux se han analizado según se indica en la sección de Materiales y Métodos (apartado 2.8). En el péptido YI12, la magnitud del valor promedio $\Delta\delta^{\text{Av}}$, que representa el efecto de la unión, es pequeña para todos los residuos con un valor medio de 0.031 ppm, figura 3.15.

Desde el punto de vista metodológico se procedió a examinar la interacción de todos los péptidos diseñados, tabla 3.2, con Fondaparinux en las mismas condiciones descritas para el péptido cabeza de serie.

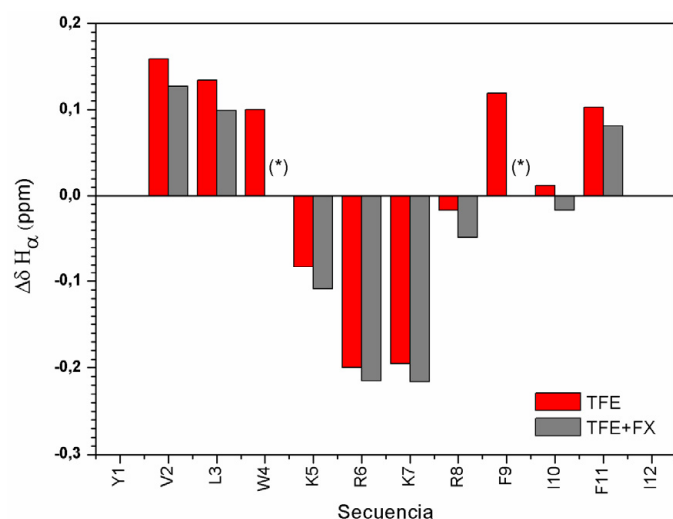


Figura 3.13. Perfiles de $\Delta\delta_{H\alpha}$ en presencia de Fondaparinux y TFE en gris y con TFE en rojo del péptido YI12. Los símbolos (*) indican que esos valores no se han podido determinar.

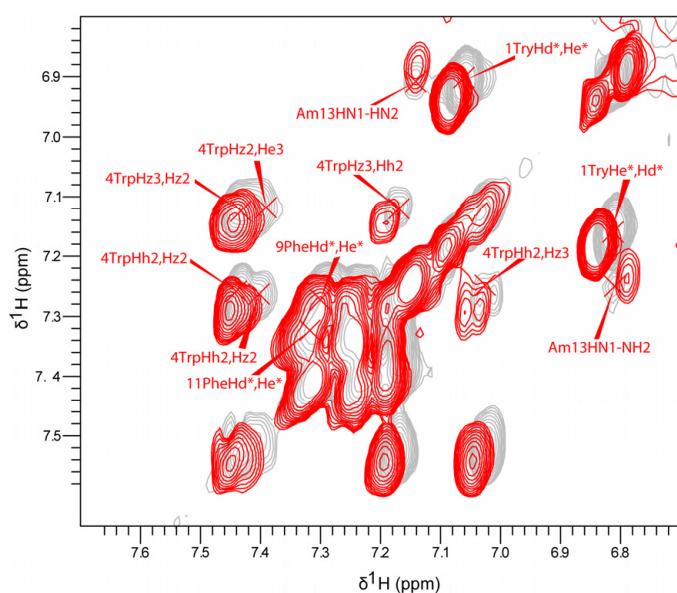


Figura 3.14. Región de los 1H aromáticos de los experimentos 1H - 1H TOCSY del péptido YI12 con TFE en rojo y con TFE y Fondaparinux en gris.

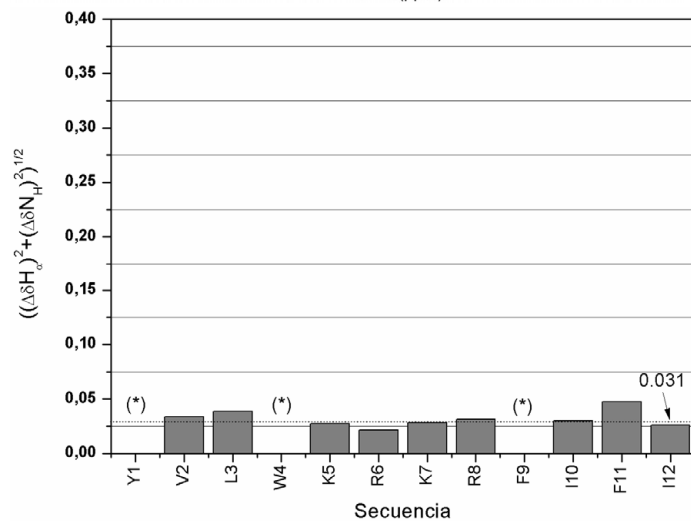


Figura 3.15. Cartografiado de desplazamientos químicos del péptido YI12 en su interacción con Fondaparinux en TFE. Los símbolos (*) indican que los valores de los desplazamientos químicos de los $H\alpha$ o de los H_N no se han podido determinar. La línea punteada indica el promedio de $\Delta\delta^{Av}$.

3.5. Interacciones de los péptidos con residuos catiónicos centrales mutados con Fondaparinux

En el péptido YI12Q5, en el que se varía el residuo catiónico de Lys en la posición 5 por una Gln que no tiene carga, es difícil analizar la interacción con Fondaparinux a partir de los desplazamientos químicos debido a que las señales en los espectros de RMN son, en general, tan anchas que fue imposible conseguir la asignación completa, figura 3.16. En concreto, no se obtuvieron los valores para los H_α de las posiciones 2, 3, 4, 5, 7 y 9. Este ensanchamiento impide la caracterización en detalle. Sin embargo, indica que la interacción entre el péptido y el Fondaparinux se produce, a pesar de que no se mantiene el motivo CPC' en la secuencia. Es interesante señalar que este péptido contiene el fragmento XBBBX que forma parte de la secuencia de tipo XBBBXBBBXBBX, de unión a heparina (Sobel et al. 1992), la cual podría jugar un papel importante en la unión con el polisacárido de Fondaparinux.

En el caso del péptido YI12Q6, ocurre algo parecido aunque menos acusado que en el caso anterior. Las señales H_α correspondientes al giro central son anchas (residuos 4, 6 y 7). A pesar de ello se adivina una tendencia similar a la del péptido YI12. En YI12Q6 el motivo CPC' se conserva y si nos fijamos en la secuencia leída del extremo C-terminal al N-terminal (C-terminal IFIFRKQKWL VY N-terminal), contiene el motivo XBBXBX presente en los sitios de unión más frecuentes a heparina (Cardin & Weintraub 1989). Se observan cambios en los desplazamientos químicos debidos a la unión de Fondaparinux especialmente en Val2, Leu3 Phe9, Phe11 y protones aromáticos. El ensanchamiento de señales y los cambios en los desplazamientos químicos apuntan a que el péptido YI12Q6 interacciona con Fondaparinux, lo que puede deberse tanto a la presencia del motivo CPC' conservado, como del motivo XBBXBX (Cardin & Weintraub 1989). La presencia de estos motivos de unión a heparina podría compensar una posible interacción más débil causada por la sustitución de Arg, ya que estudios previos han demostrado que los residuos de Arg forman fuertes enlaces de hidrógeno con la heparina (Fromm et al. 1995).

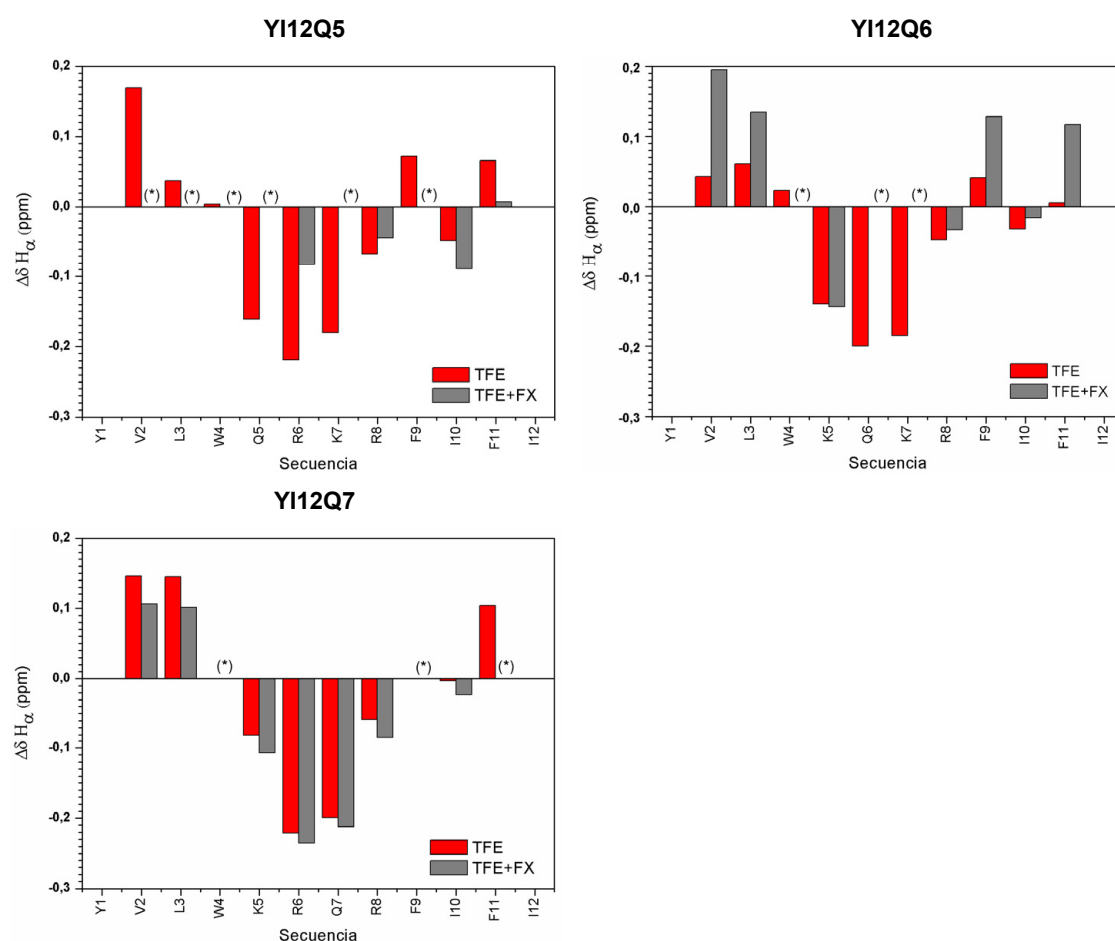


Figura 3.16. $\Delta\delta$ de H_α para los péptidos YI12Q5, YI12Q6 y YI12Q7 en presencia de Fondaparinux y TFE y TFE. El péptido YI12Q8 no se muestra porque no ha sido posible asignar el desplazamiento químico de ninguno de los protones H_α en presencia de Fondaparinux y TFE. Los símbolos (*) indican los valores que no han podido ser determinados.

En el péptido YI12Q7 la sustitución del residuo catiónico Lys de la posición 7 por una Gln, hace que contenga la secuencia de reconocimiento de interacción con heparina XBBXB (Cardin & Weintraub 1989) además de conservar el motivo CPC'. La mala calidad de los espectros hace inobservables las señales de H_α en los residuos en las posiciones 4, 9, 11, que están cerca del agua. De todas formas, la similitud del perfil de $\Delta\delta$ de este péptido al de YI12, así como las variaciones en la posición de todos los protones de los anillos aromáticos del péptido debidas a la presencia de Fondaparinux, evidencian la interacción entre ambos.

En los espectros adquiridos con el péptido YI12Q8 en presencia de Fondaparinux y TFE sólo se observan algunas señales correspondientes a los átomos de las cadenas laterales. El ensanchamiento debido a la interacción afecta a todos los H_N del esqueleto peptídico. Además de este hecho, la interacción entre Fondaparinux y el péptido YI12Q8, en presencia

Y112Q5

$\delta^1\text{H}$ (ppm)

$\delta^1\text{H}$ (ppm)

Y112Q6

$\delta^1\text{H}$ (ppm)

$\delta^1\text{H}$ (ppm)

Y112Q7

$\delta^1\text{H}$ (ppm)

$\delta^1\text{H}$ (ppm)

Si bien el análisis de la interacción péptido/Fondaparinux a partir de los núcleos del esqueleto presenta problemas debido a los ensanchamientos comentados, este efecto es menos evidente en las señales correspondientes a los anillos aromáticos de Tyr1, Trp4 y Phe9. Esto es una ventaja que nos permite monitorizar los cambios en el entorno de dichos anillos que se dan a raíz de la interacción. Las diferencias son del mismo orden en los tres péptidos YI12F6Q5, YI12F6Q6 y YI12F6Q7, aunque los sentidos de los cambios (hacia campo alto o hacia campo bajo) son específicos de cada interacción. Así, por ejemplo, el protón Hz2 del anillo indólico del Trp4 sufre cambios (complejo/libre) de +0.06 ppm en

YI12F6Q5, de -0.09 ppm en YI12F6Q6 y de -0.02 ppm en YI12F6Q7, figura 3.17. Esto nos indica que los entornos químicos de estos núcleos están afectados por la unión a Fondaparinux.

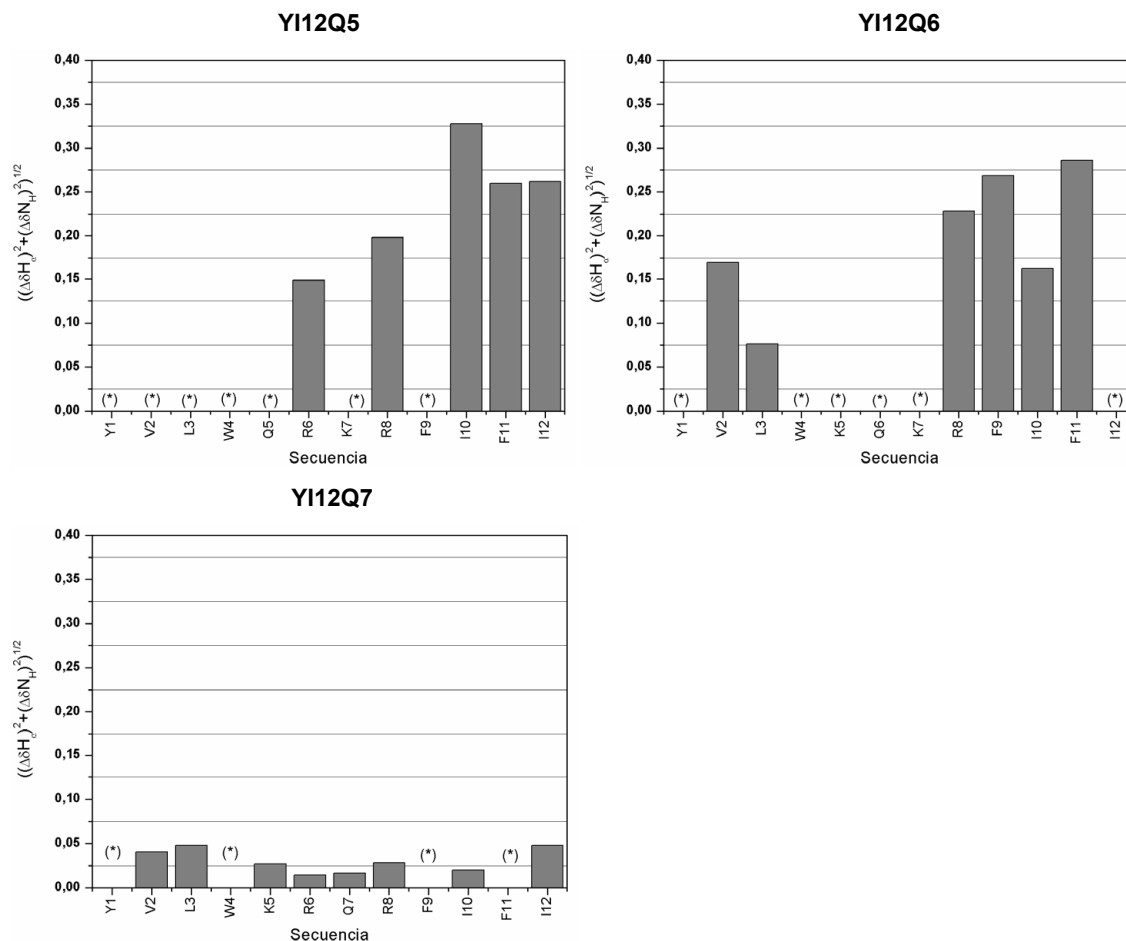


Figura 3.18. Cartografiado de desplazamientos químicos de los péptidos YI12Q5, YI12Q6 EYI12Q7. Los (*) indican valores que no han sido determinados.

Finalmente, para cuantificar de una manera objetiva el efecto de Fondaparinux sobre el esqueleto de los diferentes péptidos, se ha procedido tal y como se ha hecho con el péptido cabeza de serie, figura 3.15. En la figura 3.18 se han representado los valores absolutos de las variaciones debidas a la interacción molecular. Es evidente que las variaciones observadas para YI12Q5 y YI12Q6 son del mismo orden mientras que el efecto causado por Fondaparinux sobre YI12Q7 es mucho menor, en el rango del observado para YI12, figura 3.15.

Todos estos resultados indican que la mutación en la posición 7, no modifica la conformación peptídica presente en Y112. Esta conformación debe ser ya óptima para la interacción por lo que se observan muy pocos cambios con respecto al péptido sin Fondaparinux. Otra cuestión es lo que sucede con las mutaciones de las posiciones 5, posición C del motivo CPC', o 6, adyacente a la misma. Las variaciones son mayores en magnitud, lo que sugiere que la estructura de estos péptidos no es adecuada para la unión y precisa de una reordenación estructural para que esta pueda llevarse a cabo. La unión con el ligando induce este cambio en la geometría de la forma unida con respecto a la del péptido libre.

3.6. Interacciones de los péptidos mutados en los extremos con Fondaparinux

La sustitución de la Tyr del extremo N-terminal, que es polar, por una Phe, que es no polar, conduce a cambios muy significativos en cuanto a la tendencia general de los $\Delta\delta$ cuando FI12 se une a Fondaparinux. Una vez más las señales H_N-H_α de los residuos Trp4 y Phe9 no se han podido asignar. Hay que señalar que la magnitud de las perturbaciones con respecto al random coil en presencia del derivado de heparina, Fondaparinux, figura 3.19, son menores que los mostrados por el péptido en ausencia de ligando y menores que las que presenta el péptido cabeza de serie, figura 3.13.

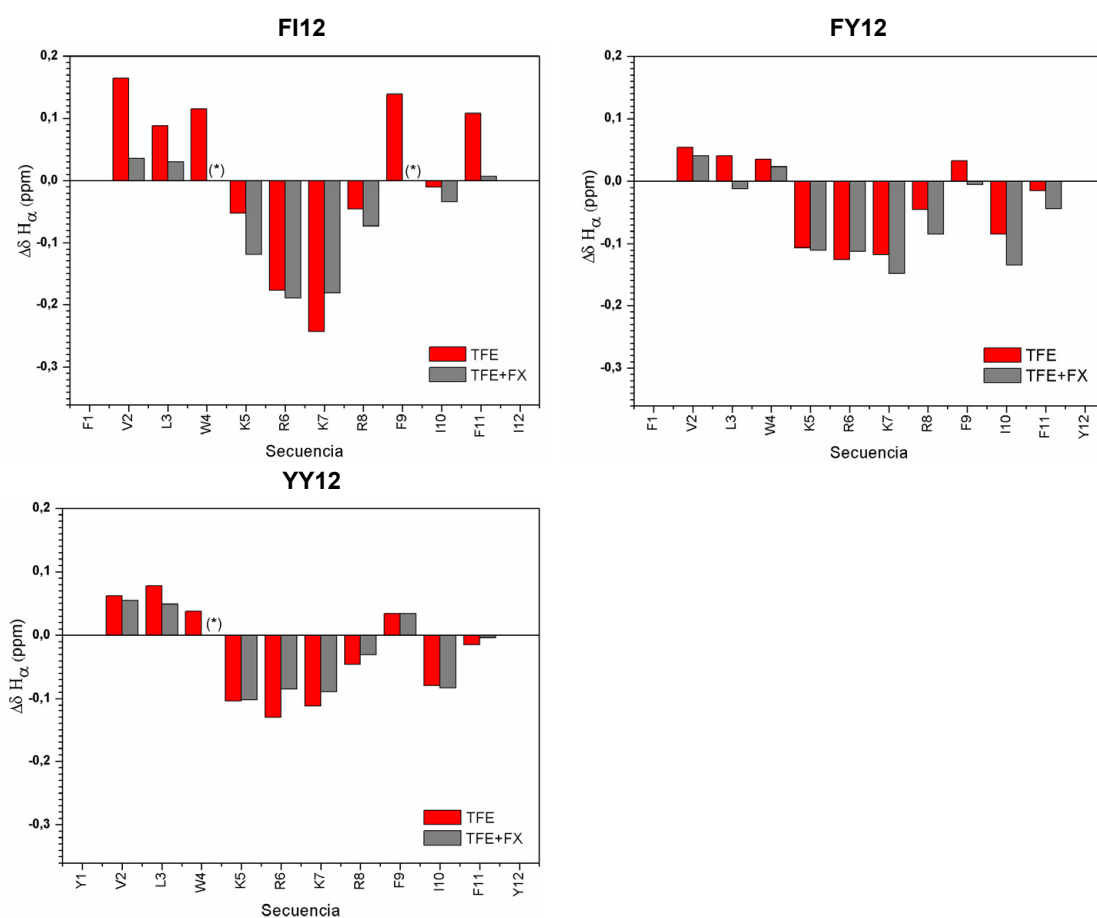


Figura 3.19. Perfiles de $\Delta\delta_{H_\alpha}$ de los péptidos FI12, FY12 e YY12 en presencia de Fondaparinux y TFE en gris y en TFE en rojo ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm). Los símbolos (*) indican valores de H_α que no han podido ser determinados.

Esta diferencia en el comportamiento del péptido FI12 cuando interacciona con Fondaparinux en relación al del péptido cabeza de serie YI12, especialmente en el extremo N-terminal, puede explicarse por el cambio de la cadena lateral de Tyr a Phe, donde el grupo OH presente en la Tyr debe jugar un papel importante en la interacción y en el mantenimiento

[illegible]

En el péptido FY12, además de modificarse la posición 1 se sustituye el residuo Ile12 del extremo C-terminal por una Tyr. Observamos muy pocos cambios en el esqueleto en presencia o no de Fondaparinux, y el ensanchamiento de señales no es muy importante habiéndose podido asignar los H_α de todos los amino ácidos de la secuencia. Es interesante mencionar que si bien hay variaciones significativas en las posiciones de los protones aromáticos de Trp4, Phe9 y Phe11, las señales del anillo aromático de la posición 12 no sufren variaciones ($\Delta\delta$ complejo/libre <0.003 ppm) por lo que este anillo no debe participar en la unión según se deduce de los espectros de RMN, figura 3.20.

El péptido YY12 presenta, en general, cambios de $\Delta\delta$ de menor magnitud que FY12, manteniéndose un perfil similar antes y después de la interacción. La diferencia más notable es que en este caso se produce un ensanchamiento muy importante de las señales pertenecientes a los protones de todos los grupos aromáticos, figura 3.20.

La interacción de los péptidos con mutaciones en los extremos está representada de forma cuantitativa en la figura 3.21. Es evidente que los cambios aparecen a lo largo de toda la secuencia y no localizados en una de sus partes, en una magnitud superior a la que presenta el cabeza de serie Y112 (media 0.031 ppm, figura 3.15). Estas variaciones unidas a los cambios observados en la región aromática permiten confirmar la interacción y evaluar de manera cualitativa la reorganización estructural que el ligando produce en el péptido en presencia de TFE.

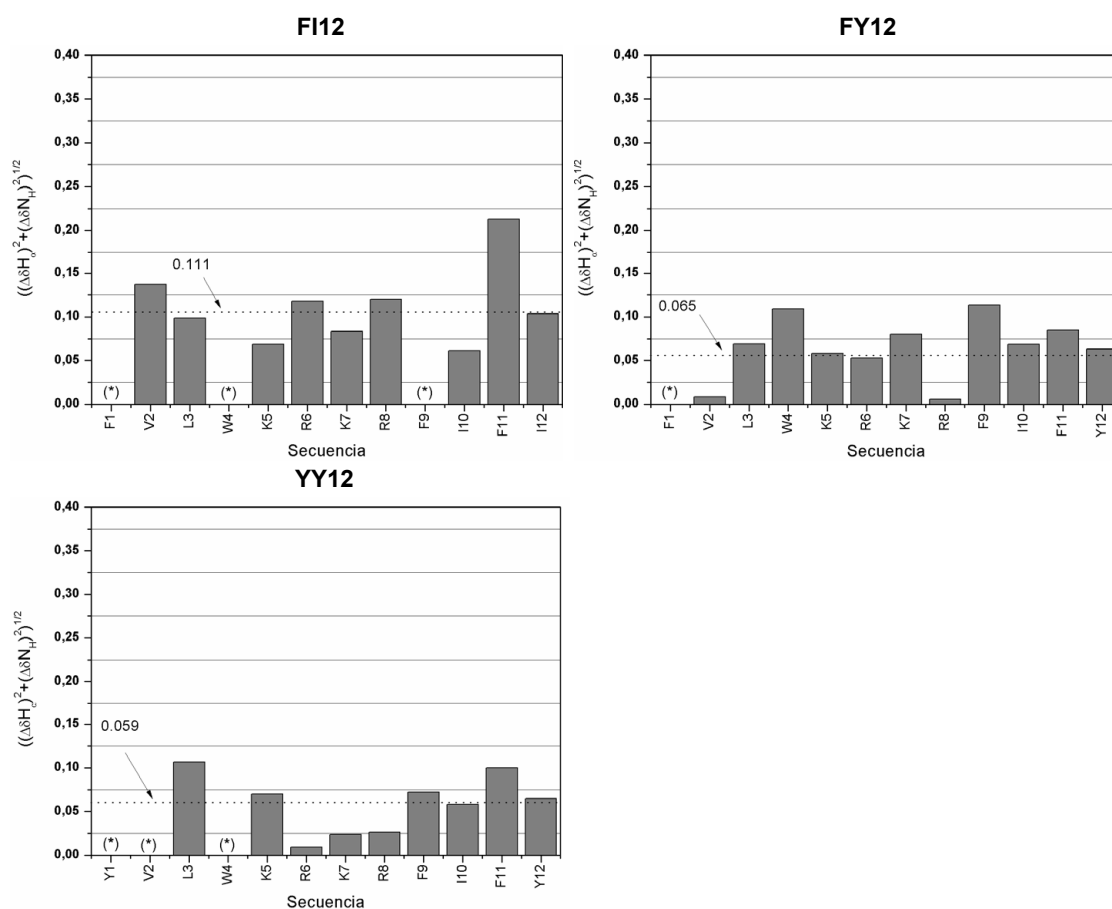


Figura 3.21. Cartografiado de desplazamientos químicos de los péptidos FI12, FY12 e YY12 donde el promedio se indica con una línea discontinua. Los símbolos (*) muestran valores que no se han podido determinar.

3.7. Interacciones del péptido YI12M9 con Fondaparinux

Este es el péptido que fue originalmente diseñado para unirse a LPS y contiene el motivo CPC' en los residuos Lys5-Tyr1-Arg8. Bhattacharjya y colaboradores (Bhattacharjya et al. 2007), determinaron su estructura en presencia de LPS y propusieron que los cuatro residuos del centro catiónico interaccionaban con la cargas negativas presentes en la superficie de los LPS, figura 3.6, que en nuestro caso correspondería a las interacciones con los grupos sulfatos que tienen carga negativa del pentasacárido Fondaparinux. Es interesante observar como la sustitución de la Phe en posición nueve por una Met tiene como resultado que las señales del espectro se mantienen finas y que los protones de los anillos aromáticos (Tyr1, Trp4 y Phe11) permanecen prácticamente inalterados en presencia de Fondaparinux, lo que apunta a que la interacción no se produce en este caso o no se mantiene en presencia de TFE.

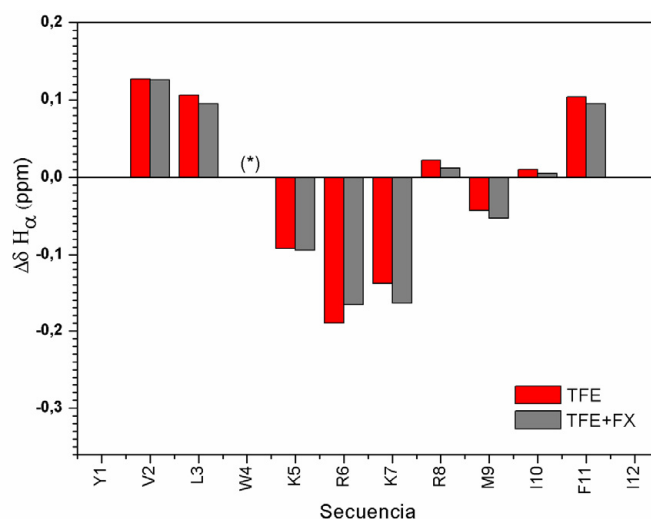


Figura 3.22. Desplazamientos químicos conformacionales del péptido YI12M9. El símbolo (*) indica los valores que no han podido ser determinados.

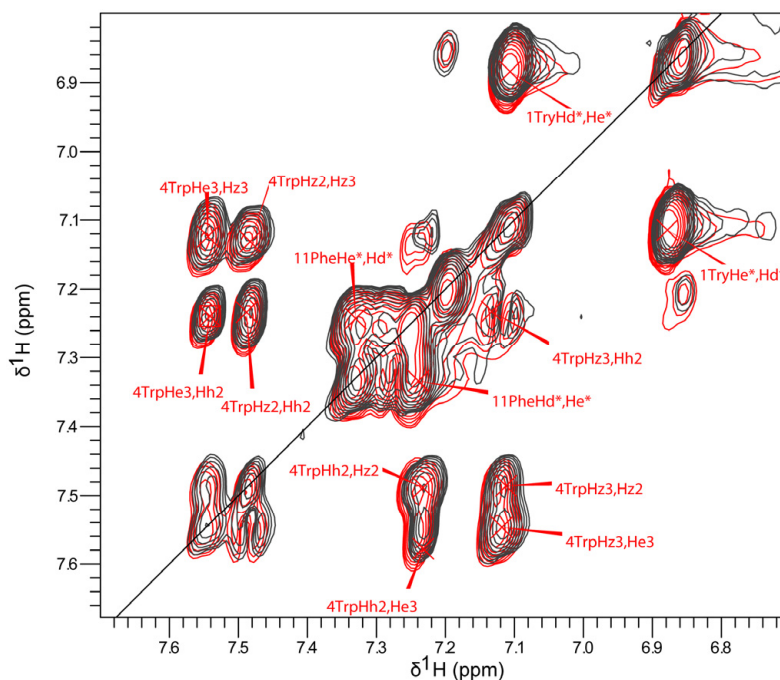


Figura 3.23. Región de los ^1H aromáticos de los experimentos TOCSY ^1H - ^1H del péptido Y112M9, en rojo en TFE y en gris en TFE y Fondaparinux.

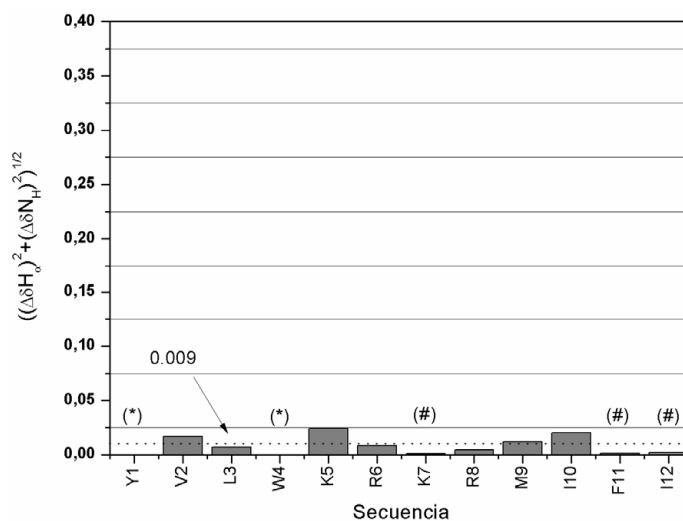


Figura 3.24. Cartografiado de desplazamientos químicos del péptido Y112M9 donde el promedio se indica con una línea discontinua. Los símbolos (*) muestran valores que no se han podido determinar y los (#) valores muy próximos a cero.

4. DISCUSIÓN

El diseño de péptidos sintéticos y la determinación de la estructura tridimensional de los complejos LPS/péptido podría proporcionar información valiosa para el desarrollo de nuevos compuestos antimicrobianos. Sin embargo, la determinación de la estructura a resolución atómica de estos complejos continúa siendo un problema de difícil solución debido a la agregación y precipitación de los péptidos en presencia de LPS (Pristovsek & Kidric 2004).

Aunque se conocía de antemano que la estructura que adoptan dos de los péptidos estudiados en este capítulo, concretamente de YI12M9 y YI12, es de tipo β -boomerang en presencia de LPS, y probablemente lo sea en heparina, en nuestros ensayos no hemos encontrado las condiciones adecuadas en las cuales se pudiesen determinar por RMN las estructuras que nos permitiesen realizar un análisis detallado de las interacciones que se dan entre la heparina y los péptidos. A pesar de las dificultades y en un intento de hacer una caracterización a nivel atómico de estas interacciones, en este capítulo hemos analizado el reconocimiento de un derivado del glicosaminoglicano heparina, el Fondaparinux, por péptidos que mimetizan proteínas que contienen el motivo clip CPC' (catiónico-polar-catiónico). Para ello se han utilizado como modelo péptidos tipo boomerang y los desplazamientos químicos de protón en diversas condiciones.

4.1. Estructura de los péptidos en agua y en la mezcla agua/TFE

Para hacer un análisis completo de las interacciones, hay que empezar por saber y comprender como es la estructura de las moléculas en su estado libre. Así, a partir de los datos de RMN, es interesante comprobar cómo cambia la estructura de los péptidos al cambiar el medio de una solución acuosa a la que contiene un porcentaje considerable (cerca del 30%) de TFE. Es conocido que la RMN es una técnica muy útil y utilizada para caracterizar conformaciones y tendencias conformacionales en péptidos utilizando los desplazamientos conformacionales de protón. Así, podemos decir que unos residuos están en conformación de hebra β cuando los $\Delta\delta$ de los H_α son positivos, que forman un giro o una hélice si éstos son negativos, y que no presentan ninguna tendencia cuando los valores se acercan a 0 (Wishart et al. 1992).

Como ya se ha expuesto en la Introducción de este capítulo, el péptido YI12, que tomamos aquí como referencia, ha sido estudiado anteriormente por RMN y se ha demostrado que es desestructurado en solución acuosa pero que adopta una estructura de β -boomerang en presencia de micelas de LPS (Bhunja et al. 2009). Esta estructura está

caracterizada por presentar valores de desplazamientos conformacionales de protón positivos en los extremos y negativos en el centro de la secuencia,

Todos los péptidos estudiados en este trabajo son muy flexibles en agua, sobre todo en los extremos donde no se observa una tendencia estructural clara. Sin embargo, todos muestran una marcada tendencia a la formación de un giro en la parte central de la secuencia donde se encuentran los residuos catiónicos, Lys5, Arg6, Lys7 y Arg8, en el péptido cabeza de serie. Este rasgo estructural es independiente de las mutaciones propias de los diferentes péptidos.

La presencia de TFE, un solvente que se caracteriza por estabilizar las estructuras secundarias de los péptidos, parece estabilizar la estructura de β -boomerang en todos los péptidos, salvo en el caso de YI12Q6 e YI12Q8. Observando que estos dos péptidos presentan mutaciones en las que se eliminan los residuos cargados de Arg en las posiciones 6 (YI12Q6) y 8 (YI12Q8), podemos deducir que las interacciones que pueden derivarse de la zona catiónica son importantes para inducir la correcta estructuración de las hebras N- y C-terminales en TFE. En este medio, los péptidos YI12Q6 y YI12Q8 mantienen el giro de los residuos centrales pero los extremos de la cadena están en una situación desordenada más parecida a la que poseen en agua. Por ello podemos decir que no adoptan la estructura esperada de β -boomerang. Sorprendentemente, la sustitución de los residuos de Lys en posiciones 5, en el péptido YI12Q5, y 7, en YI12Q7, no impide la formación de la estructura de β -boomerang en TFE. Este resultado sugiere, en primer lugar, que son los contactos específicos de determinados residuos catiónicos, concretamente los de las posiciones 6 y 8, los que resultan determinantes en la estructuración del péptido en TFE y, en segundo lugar, que las interacciones de los residuos de Arg son decisivas para la estructuración de los péptidos ya que su eliminación supone una cierta “desestructuración”.

Por otra parte, los datos de FY12 e YY12 indican que probablemente la estructura de β -boomerang se forme aunque en menor medida que en los otros casos. Es decir, la inclusión de un residuo aromático polar en el extremo C-terminal de la secuencia, no ayuda a la estabilización de la estructura de β -boomerang, lo que sugiere que un aminoácido hidrofóbico como Ile es el ideal para promover este tipo de plegamiento en TFE.

En TFE, la estructura adoptada por los péptidos de este estudio es similar a la que se puede favorecer en presencia de micelas. Existen antecedentes en la bibliografía que avalan esta suposición. Concretamente YI12 adquiere en LPS (Bhattacharjya et al. 2007) la misma estructura de β -boomerang que se ha caracterizado aquí para varios de sus mutantes en TFE. Por ello, la estructura adoptada en este entorno hidrofóbico debe ser la predeterminada

para las interacciones con el derivado de heparina Fondaparinux y es probable que los cambios que se produzcan por la interacción sean mínimos en el entorno de TFE cuando la estructura sea cercana a la que posee el péptido en el complejo. Por otro lado, nuestros datos indican que el balance de cargas y polaridades es muy crítico para este tipo de estructura en este medio, por lo que sutiles diferencias en las interacciones que puedan darse con los ligandos, pueden tener consecuencias importantes en la adopción de la estructura competente para la unión.

Una vez dilucidada la estructura de partida, vamos a discutir los cambios que se han observado con Fondaparinux.

4.2. Interacciones entre los péptidos y Fondaparinux en presencia de TFE.

El análisis de las gráficas de la figuras 3.15, 3.18, 3.21 y 3.24 nos permite observar que los desplazamientos químicos de los protones del esqueleto de algunos de los péptidos (concretamente los péptidos YI12, YI12Q7, FY12 y YI12M9) apenas se afectan por la presencia de Fondaparinux, ya que muestran unos valores de $\Delta\delta^{Av}$ muy bajos que pueden explicarse por dos situaciones diferentes. Una de ellas es que, tal como se ha comentado anteriormente, la conformación adoptada por la cadena principal en presencia de TFE está ya adaptada a la necesaria para la unión. Teniendo en cuenta esta premisa y que el reconocimiento de Fondaparinux tiene que ser forzosamente superficial dada la naturaleza de las moléculas participantes (Langeslay et al. 2012), en este supuesto el péptido sólo precisaría adaptar las cadenas laterales para optimizar la interacción. El otro escenario posible contempla que los bajos valores $\Delta\delta^{Av}$ de los núcleos del esqueleto provengan de interacciones muy débiles o inclusive de la falta de interacción con el Fondaparinux. En este caso las cadenas laterales no serán sensibles a la presencia del ligando y mantendrán el comportamiento del péptido libre en TFE.

Así pues, de los datos analizados en este capítulo se deduce que hay péptidos en las dos situaciones. Los resultados de RMN apuntan a que YI12M9 podría estar en el segundo escenario, caracterizado por señales finas en los espectros de RMN, figura 3.23, en todas las condiciones y ausencia de cambios en las posiciones de los protones del esqueleto y aromáticos después de la mezcla con Fondaparinux. Esto indica claramente, la necesidad de la presencia del residuo aromático Phe en la posición 9 para la interacción con Fondaparinux, al menos en presencia de TFE. Este resultado va en la línea de resultados que se ha descrito en la literatura señalando la importancia de este tipo de residuos en la unión a heparina y otros derivados de glicosaminoglicanos (García-Mayoral et al. 2010). Es importante resaltar que la mutación de Phe por Met en la posición 9 no anula la capacidad del péptido de formar

la estructura de β -boomerang aunque como ya se ha expuesto en los resultados de la sección 3.7, el perfil de $\Delta\delta$ sugiere que la geometría del giro debe ser ligeramente diferente a la que presentan sus homólogos.

Por otro lado, YI12Q7, en el que se ha eliminado la carga de la posición 7, es el que presenta unos parámetros de RMN más parecidos a YI12. En la interacción de YI12Q7 con Fondaparinux se producen los mismos cambios en el esqueleto y en las señales de los anillos aromáticos Tyr1, Trp4 y Phe9 que en YI12. Esto añade más importancia a la idea expuesta anteriormente de que la carga en esta posición no determina el establecimiento de la estructura de β -boomerang y además no es imprescindible para el reconocimiento del ligando. En esta misma línea se puede concluir que las cargas de las posiciones 5 y 6, por si mismas tampoco son esenciales para interactuar con Fondaparinux. Tanto YI12Q5 como YI12Q6 interactúan con Fondaparinux, sufriendo, al menos en YI12Q6 de manera evidente, un cambio conformacional en TFE que favorece claramente la estabilización de la estructura de β -boomerang que no estaba poblada en ausencia de ligando. Así pues, de estos datos se deduce que la carga de los residuos catiónicos frente a residuos polares neutros en las posiciones 5 y 6 es esencial para mantener la estructura competente en ausencia de ligando, pero no es imprescindible la presencia de todas ellas para la unión a Fondaparinux en TFE. Esto está de acuerdo con los resultados descritos por Fromm y colaboradores (Fromm et al. 1997) que indican que un GAG altamente sulfatado, como la heparina, interactúa más fuertemente con las secuencias que contienen un sitio de unión complementario de alta densidad de carga positiva; mientras que el sulfato de heparán, con menos grupos cargados negativamente y más espaciados (como el Fondaparinux), interactúa mejor con una secuencia que tenga más separados los residuos catiónicos.

La sustitución del residuo polar de Tyr de la posición 1, por un residuo de Phe, en el péptido FI12 no muestra apenas diferencias con respecto al péptido cabeza de serie YI12 cuando está en agua, en TFE o en TFE y Fondaparinux. A pesar de que este péptido adopta la estructura de β -boomerang en TFE y en TFE y Fondaparinux, la diferencia de los valores de desplazamientos químicos indican que en presencia de Fondaparinux se desestabiliza la estructura (véase la diferencia entre las barras rojas y grises, figura 3.19) tanto de la hebra N-terminal como la C-terminal. Estos resultados indican que la Tyr de la posición 1 participa directamente en la interacción y que su sustitución conlleva una pérdida de estructura.

El cambio del residuo polar del extremo N-terminal al C-terminal, en el péptido FY12, evidencia que la estructura de β -boomerang ya estaba formada en presencia de TFE. Sólo se observan diferencias en los residuos de Phe9, resultado de la presencia del residuo polar en

la posición 12, y en Leu3, probablemente como consecuencia de las interacciones con Trp4 (cuyas interacciones con Fondaparinux se observan con claridad si nos fijamos en los protones aromáticos del Trp4, figura 3.20).

Las diferencias entre el comportamiento de FY12 e YY12 son mínimas a pesar de la adición de un segundo residuo polar en este último caso.

4.3. Conservación del motivo grapa CPC'

El motivo grapa, catiónico-polar-catiónico, CPC', corresponde estrictamente en estos péptidos con Lys5 (C), Tyr1 (P) y Arg8 (C'). Los péptidos estudiados que conservan dicho motivo YI12, YI12M9, YI12Q6, YI12Q7 y YY12, muestran distinto comportamiento. Todos ellos son capaces de formar la estructura de β -boomerang, sin embargo no todos unen Fondaparinux en TFE. Los parámetros de RMN han demostrado que la presencia del residuo aromático en posición 9 es clave, ya que su pérdida en el caso de YI12M9 impide la interacción. Luego, la grapa CPC' es condición necesaria pero no suficiente para la unión.

5. CONCLUSIONES

El objetivo inicialmente propuesto de resolver la estructura del péptido YI12 y sus derivados cuando interaccionan con heparina y/o con el análogo Fondaparinux, con el fin de determinar la importancia del motivo CPC' en la interacción, no ha podido alcanzarse. El motivo principal han sido las dificultades experimentales que se han detallado a lo largo del capítulo y que han hecho imposible la pretendida resolución estructural. Sin embargo, a pesar de ello, el análisis de los parámetros de RMN, fundamentalmente los desplazamientos químicos y las anchuras de línea, nos permiten extraer las siguientes conclusiones:

Tanto el péptido original descrito por Bhattacharjya, YI12M9, como YI12 y sus derivados, presentan una estructura en solución acuosa caracterizada por una cierta propensión a la formación de un giro en los residuos catiónicos centrales, con los extremos N- y C-terminales desordenados. Esta tendencia conformacional se manifiesta en perfiles de los desplazamientos conformacionales $\Delta\delta$ de H_α con valores más negativos en la región del giro.

La presencia de TFE induce la estabilización de la estructura secundaria de β -boomerang en todos los péptidos excepto en YI12Q6 y YI12Q8. Este resultado indica que los residuos Arg6 y Arg8 son determinantes para la formación de la estructura β -boomerang en presencia de TFE. Esto podría deberse a la participación de estos residuos de Arg en interacciones de tipo π -catión, en concreto, Arg6/Phe9 y Arg8/Phe11.

La estructura global de β -boomerang que adoptan los péptidos en TFE se mantiene en presencia de Fondaparinux y TFE. Las variaciones estructurales debidas a la interacción, monitorizadas a partir de los datos de los desplazamientos químicos, son pequeñas. Esto indica que, en presencia de TFE, los péptidos adoptan una estructura óptima para la interacción que sólo debe ajustar la disposición de las cadenas laterales cuando interacciona con Fondaparinux.

Los péptidos YI12Q5, YI12Q6, YI12Q7 e YI12Q8, aunque con diferencias, interaccionan con Fondaparinux en presencia de TFE, lo cual se refleja principalmente en las variaciones de los desplazamientos químicos, en el ensanchamiento de las señales y la formación de un precipitado en el caso de YI12Q8. Así pues, ninguno de los residuos catiónicos de las posiciones 5, 6, 7 y 8, por sí mismos, resultan indispensables para la interacción con Fondaparinux. Esto es interesante puesto que se ha propuesto que las interacciones con los grupos sulfato del Fondaparinux se realizan a través de los residuos catiónicos centrales.

Nuestros resultados sugieren que la interacción requiere la existencia de una cierta densidad de carga positiva, no siendo imprescindible la presencia puntual de un determinado residuo.

La presencia de residuos de Tyr, ya sea en el extremo N-terminal (YY12) o en el C-terminal (FY12), parece jugar un papel importante en las interacciones con el ligando. La ausencia de la Tyr en el péptido FI12, conduce a una desestabilización de la estructura de β -boomerang cuando interacciona con Fondaparinux en presencia de TFE. Por tanto, la presencia de una Tyr en los extremos del péptido es un requerimiento para mantener la estructura del complejo con Fondaparinux.

La identidad de los desplazamientos químicos de los átomos del esqueleto y de las cadenas laterales del péptido YI12M9, cuando está en TFE o en TFE y Fondaparinux, sugiere que no hay interacción con Fondaparinux, o que no ha sido posible su monitorización en las condiciones utilizadas en este trabajo. Este resultado difiere del publicado acerca de la interacción observada entre YI12M9 y LPS. Por tanto, a pesar de la analogía entre los LPS y el derivado de heparina Fondaparinux en TFE que hemos utilizado en nuestro estudio, las interacciones que se establecen entre el péptido y los dos ligandos no son iguales en las condiciones exploradas.

La presencia en los péptidos de las secuencias de reconocimiento de heparina del tipo XBBBX y XBBXBX puede contribuir a su interacción con Fondaparinux, aunque los resultados obtenidos no son suficientes para dilucidar su importancia relativa.

Los parámetros de RMN no muestran un patrón diferencial claro entre los péptidos que conservan el motivo CPC', Lys5-Tyr1-Arg8, (YI12, YI12Q6, YI12Q7, YY12 e YI12M9) y los que no lo conservan (YI12Q5, YI12Q8, FI12 y FY12). Esto indica que las interacciones en TFE con el derivado de heparina, Fondaparinux, son independientes de la conservación de dicho motivo.

6. BIBLIOGRAFÍA

- Aue WP, Bartholdi D, Ernst RR (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. . Journal of Chemical Physics 64, 2229-2246.
- Bax A, Davis DG (1985) Practical aspects of two-dimensional transverse NOE spectroscopy. Journal of Magnetic Resonance 63, 207-213.
- Bhattacharjya S, Domadia PN, Bhunia A et al. (2007) High-resolution solution structure of a designed peptide bound to lipopolysaccharide: transferred nuclear Overhauser effects, micelle selectivity, and anti-endotoxic activity. Biochemistry 46, 5864-5874.
- Bhunia A, Mohanram H, Domadia PN et al. (2009) Designed beta-boomerang antiendotoxic and antimicrobial peptides: structures and activities in lipopolysaccharide. The Journal of biological chemistry 284, 21991-22004.
- Bishop CM, Walkenhorst WF, Wimley WC (2001) Folding of beta-sheets in membranes: specificity and promiscuity in peptide model systems. Journal of molecular biology 309, 975-988.
- Braunschweiler L, Ernst RR (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. Journal of Magnetic Resonance 53, 521-528.
- Caldwell EE, Nadkarni VD, Fromm JR et al. (1996) Importance of specific amino acids in protein binding sites for heparin and heparan sulfate. The international journal of biochemistry & cell biology 28, 203-216.
- Capila I, Linhardt RJ (2002) Heparin-protein interactions. Angew Chem Int Ed Engl 41, 391-412.
- Cardin AD, Weintraub HJ (1989) Molecular modeling of protein-glycosaminoglycan interactions. Arteriosclerosis 9, 21-32.
- Casu B, Guerrini M, Torri G (2004) Structural and conformational aspects of the anticoagulant and anti-thrombotic activity of heparin and dermatan sulfate. Current pharmaceutical design 10, 939-949.
- Clore GM, Gronenborn AM (1982) Theory of the time dependent transferred nuclear Overhauser effect: applications to ligand protein complexes in solution. Journal of Magnetic Resonance 48, 402-417.
- Choay J, Lormeau JC, Petitou M (1981) [Low molecular weight oligosaccharides active in plasma against factor Xa (author's transl)]. Annales pharmaceutiques francaises 39, 37-44.
- Duerst R, Verbach A (1965) Accurate NMR temperature Measurements. Review of Scientific Instruments 36, 1896-1897.
- Ferguson AD, Hofmann E, Coulton JW et al. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science 282, 2215-2220.
- Ferguson AD, Welte W, Hofmann E et al. (2000) A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. Structure 8, 585-592.
- Fromm JR, Hileman RE, Caldwell EE et al. (1995) Differences in the interaction of heparin with arginine and lysine and the importance of these basic amino acids in the binding of heparin to acidic fibroblast growth factor. Archives of biochemistry and biophysics 323, 279-287.
- Fromm JR, Hileman RE, Caldwell EE et al. (1997) Pattern and spacing of basic amino acids in heparin binding sites. Archives of biochemistry and biophysics 343, 92-100.
- Gandhi NS, Mancera RL (2010) Heparin/heparan sulphate-based drugs. Drug discovery today 15, 1058-1069.
- Garcia-Mayoral MF, Moussaoui M, de la Torre BG et al. (2010) NMR structural determinants of eosinophil cationic protein binding to membrane and heparin mimetics. Biophysical journal 98, 2702-2711.
- Goddard TD, Kneller DG (2005) Sparky 3, University of California, San Francisco.

- Gomes JA, Mallion RB (2001) Aromaticity and ring currents. *Chemical reviews* 101, 1349-1383.
- Han X, Bushweller JH, Cafiso DS et al. (2001) Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nature structural biology* 8, 715-720.
- Hileman RE, Fromm JR, Weiler JM et al. (1998) Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. *BioEssays : news and reviews in molecular, cellular and developmental biology* 20, 156-167.
- Jeener J, Meier BH, Bachmann P et al. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *The Journal of chemical physics* 71, 4546-4553.
- Jiménez MA, Gómez-Moreno C, Sancho Sanz J (2003) Estructura de proteína por RMN. *Ariel Ciencia*.
- Kisilevsky R, Ancsin JB, Szarek WA et al. (2007) Heparan sulfate as a therapeutic target in amyloidogenesis: prospects and possible complications. *Amyloid: the international journal of experimental and clinical investigation: the official journal of the International Society of Amyloidosis* 14, 21-32.
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* 14, 51-55, 29-32.
- Kovensky J (2009) Sulfated oligosaccharides: new targets for drug development? *Current medicinal chemistry* 16, 2338-2344.
- Kumar A, Ernst RR, Wuthrich K (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and biophysical research communications* 95, 1-6.
- Langeslay DJ, Young RP, Beni S et al. (2012) Sulfamate proton solvent exchange in heparin oligosaccharides: evidence for a persistent hydrogen bond in the antithrombin-binding pentasaccharide Arixtra. *Glycobiology* 22, 1173-1182.
- Liu D, Shriver Z, Qi Y et al. (2002) Dynamic regulation of tumor growth and metastasis by heparan sulfate glycosaminoglycans. *Seminars in thrombosis and hemostasis* 28, 67-78.
- Macura S, Ernst RR (1980) Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy. *Molecular Physics* 41, 95-117.
- Margalit H, Fischer N, Ben-Sasson SA (1993) Comparative analysis of structurally defined heparin binding sequences reveals a distinct spatial distribution of basic residues. *The Journal of biological chemistry* 268, 19228-19231.
- Peplow PV (2005) Glycosaminoglycan: a candidate to stimulate the repair of chronic wounds. *Thrombosis and haemostasis* 94, 4-16.
- Post CB (2003) Exchange-transferred NOE spectroscopy and bound ligand structure determination. *Current opinion in structural biology* 13, 581-588.
- Pristovsek P, Kidric J (2004) The search for molecular determinants of LPS inhibition by proteins and peptides. *Current topics in medicinal chemistry* 4, 1185-1201.
- Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annual review of biochemistry* 71, 635-700.
- Sibanda BL, Blundell TL, Thornton JM (1989) Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *Journal of molecular biology* 206, 759-777.
- Sobel M, Soler DF, Kermode JC et al. (1992) Localization and characterization of a heparin binding domain peptide of human von Willebrand factor. *The Journal of biological chemistry* 267, 8857-8862.
- Torrent M, Nogues MV, Andreu D et al. (2012) The "CPC clip motif": a conserved structural signature for heparin-binding proteins. *PloS one* 7, e42692.
- Turpie AG (2004) Fondaparinux: a Factor Xa inhibitor for antithrombotic therapy. *Expert opinion on pharmacotherapy* 5, 1373-1384.

- Turpie AG, Eriksson BI, Bauer KA et al. (2004) Fondaparinux. The Journal of the American Academy of Orthopaedic Surgeons 12, 371-375.
- Wishart DS, Bigam CG, Holm A et al. (1995a) ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. Journal of biomolecular NMR 5, 332.
- Wishart DS, Bigam CG, Holm A et al. (1995b) ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. Journal of biomolecular NMR 5, 67-81.
- Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31, 1647-1651.
- Wüthrich K (1986) NMR of Proteins and Nucleic Acids. J Wiley&Sons, Inc New York.
- Zasloff M (2002) Antimicrobial peptides of multicellular organisms. Nature 415, 389-395.

Capítulo IV

Estructura de péptidos que
mimetizan distintas regiones del ectodominio
de la glicoproteína gp41 del virus VIH-1

Capítulo IV:

Estructura de péptidos que mimetizan distintas regiones del ectodominio de la glicoproteína gp41 del virus VIH-1

1. INTRODUCCIÓN

1.1. El Virus de Inmunodeficiencia Humana (VIH)

El VIH produce la destrucción del sistema inmunitario de la persona infectada, por lo que es el causante del estado patológico denominado Síndrome de Inmunodeficiencia Adquirida, SIDA. Una persona padece SIDA cuando su organismo, debido a la inmunodeficiencia provocada por el VIH, no es capaz de ofrecer una respuesta inmune adecuada contra las infecciones a través de la producción de linfocitos T. Las células T son atacadas por el virus, de manera que la cantidad de las mismas desciende por debajo de 200 células por mililitro de sangre, cuando los niveles normales oscilan entre 600-1200 por mL.

Los primeros casos de SIDA se detectaron a finales de los años 70 del siglo pasado y se publicaron por primera vez en 1981, aunque existe cierta controversia acerca de los grupos que se atribuyen la primera autoría. El aislamiento y purificación del virus VIH en 1984 permitió el desarrollo de un anticuerpo, que comenzó a utilizarse para identificar a los infectados y supuso el inicio de la investigación de tratamientos para su cura. La incidencia global de la enfermedad en el año 2009 era de 33.4 millones (31.1-35.8 millones) de personas infectadas (UNAIDS 2009).

El virus VIH es un retrovirus que se clasifica dentro del género de los lentivirus. Existen dos tipos de virus VIH que se diferencian genética y antigénicamente. El VIH-1 que es el causante de la pandemia mundial, y el VIH-2 que se encuentra localizado casi exclusivamente en el oeste de África. Los dos virus presentan síntomas clínicos similares y la misma forma de transmisión, pero el desarrollo de la enfermedad del SIDA es mucho más lento en las infecciones con VIH-2 que con VIH-1.

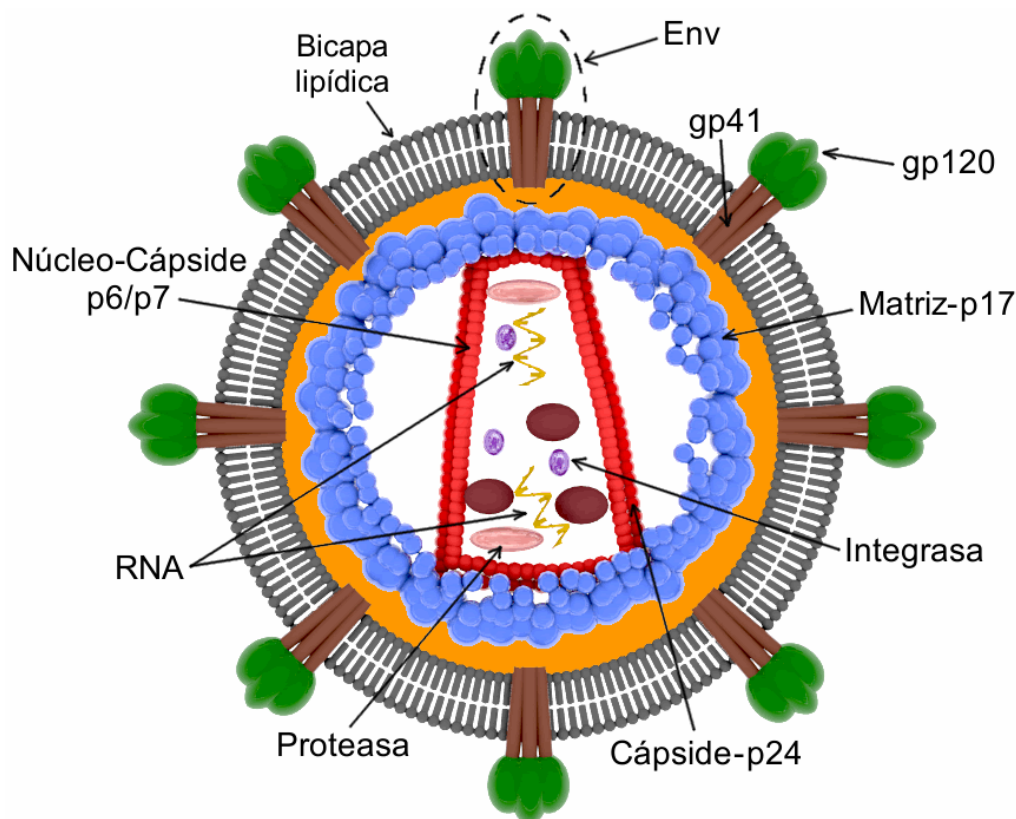


Figura 4.1. Imagen esquemática de la estructura del virión del VIH.

La partícula infectiva del VIH, denominada virión que se representa en la figura 4.1, presenta una estructura diferente a la de otros virus. Se trata de un virión prácticamente esférico de unos 120 nm de diámetro (aproximadamente 60 veces más pequeño que un glóbulo rojo). Su genoma es de 9.8 kpb (9800 pares de nucleótidos) cuando está integrado en la célula huésped. Este genoma consta de dos copias de RNA monocatenario de sentido positivo (mismo sentido que el RNA mensajero de las células huésped) que codifican para nueve genes. Las hebras de RNA se encuentran unidas a las proteínas de la núcleo-cápside (p6/p7) que, a su vez, están rodeadas por una cápside cónica compuesta por unas 2000 copias de la proteína viral p24. Dentro de la cápside se encuentran también las enzimas necesarias para el desarrollo del virión como la transcriptasa inversa, la proteasa, la ribonucleasa y la integrasa. Una matriz compuesta por la proteína viral p17 rodea la cápside

para garantizar la integridad del virión. Esta matriz, a su vez, está rodeada por la envoltura vírica que se compone de dos capas de fosfolípidos, tomados de la membrana plasmática de la célula huésped, y que incluye proteínas de la célula huésped y alrededor de 70 copias del complejo de proteínas propias del VIH conocido como Env. Los complejos Env tienen un papel relevante en el proceso de infección del virus VIH, puesto que permiten al virus adherirse y fusionarse con las dianas celulares e iniciar el ciclo infeccioso.

Estos complejos Env están formados por tres heterodímeros cuyas subunidades, las glicoproteínas gp120 y gp41, se encuentran asociadas de manera no covalente. Cada complejo Env tiene forma de “clavo” con lo que sería la “cabeza” formada por las tres glicoproteínas gp120 que sobresalen a través de la membrana vírica, y el “cuerpo” formado por las tres glicoproteínas gp41 ancladas a la membrana viral, figura 4.1. Las estructuras a baja resolución del complejo Env obtenidas mediante crio-tomografía electrónica muestran una estructura trimérica de las proteínas gp120, pero difieren en la disposición de las tres moléculas de gp41, bien en forma de tallo confirmando el modelo de “clavo”, o bien proyectándose de forma oblicua a la cabeza, en un modelo de trípode (Zanetti et al. 2006; Zhu et al. 2006).

1.2. El complejo Env: diana de anticuerpos frente a VIH

Algunos individuos infectados por el virus VIH desarrollan anticuerpos monoclonales neutralizantes (bNAbs) tras largos periodos de tiempo de infección (Karlsson Hedestam et al. 2008). Estos anticuerpos bNAbs tienen como diana el complejo Env y son capaces de bloquear la infección para una amplia gama de cepas del VIH-1 (Burton et al. 2004; Karlsson Hedestam et al. 2008). Por tanto, una terapia frente al virus VIH, alternativa a las actualmente utilizadas y diferente de la estrategia explorada en el capítulo siguiente, es el diseño de vacunas que induzcan la producción de anticuerpos cualitativamente similares a los bNAbs (Burton 2002). Los anticuerpos neutralizantes inducidos por la vacuna podrían bloquear el proceso infeccioso del VIH-1.

Se han descrito bNAbs que reconocen epítomos localizados bien en la glicoproteína gp120, o bien en la glicoproteína gp41. Entre estos últimos destacan los conocidos como 2F5, 4E10 y Z13. Los epítomos de reconocimiento en la proteína gp120 son estructuralmente muy complejos y discontinuos. En cambio, los epítomos de la proteína gp41 se localizan en la región MPER, donde se cree que existe un único epítomo lineal (Ardelt et al. 1991; Ofek et al. 2004; Cardoso et al. 2007; Pejchal et al. 2009; Zhu et al. 2011; Huang et al. 2012). Por otra parte, la proteína gp120 contiene varias regiones variables, mientras que gp41 está altamente conservada. Por ello, en general, los estudios encaminados al diseño y desarrollo de vacunas basadas en péptidos inmunogénicos se centran en la glicoproteína gp41.

1.3. Relevancia de la glicoproteína gp41 en el mecanismo de fusión de membranas

Los complejos Env juegan un papel esencial en el inicio del proceso de infección de las células hospedadoras. En primer lugar, la glicoproteína gp120, situada en la parte exterior de la superficie de la membrana en el estado nativo del virión, figura 4.2 (A), se une a los receptores celulares CD4 de los linfocitos T. Esta unión origina un cambio conformacional en el complejo Env que conduce a la formación del estado intermedio de “pre-horquilla”, figura 4.2 (B) (Chan & Kim 1998; Eckert & Kim 2001), y que produce la energía necesaria para la fusión de las membranas (Eckert & Kim 2001; Gallo et al. 2003; Melikyan 2008). En este estado de pre-horquilla la glicoproteína gp41, que es una proteína transmembrana (Roux & Taylor 2007; Karlsson Hedestam et al. 2008), se ancla a la membrana de la célula diana por la región FP (péptido de fusión) de su extremo N-terminal, mientras continúa anclada a la membrana viral por la región TMD (dominio transmembrana), figuras 4.2 (B) y 4.3. De esta forma, el virión y la célula diana se encuentran ya conectados. A continuación, esta estructura filamentosa colapsa y vuelve a plegarse sobre sí misma dando lugar a un manojo de 6 hélices (6-HB) o un trímero de horquillas, donde las tres hélices C-terminales (región CHR de gp41), figura 4.3, se empaquetan en dirección contraria al surco hidrófobo externo formado por las hélices N-terminales (región NHR de gp41), figura 4.3, (Weissenhorn et al. 1997b; Eckert & Kim 2001).

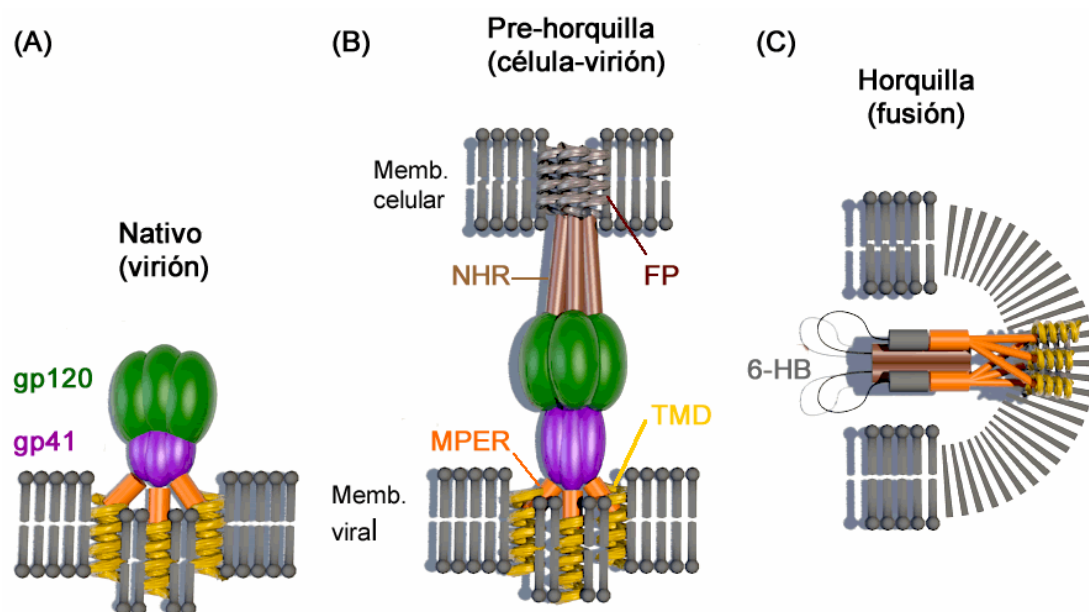


Figura 4.2. Representación esquemática de las etapas del modelo de fusión de membranas del VIH. (A) El complejo Env en el estado nativo del virión. (B) Estado intermedio de pre-horquilla, donde la región FP de gp41 se ancla a la membrana celular de la célula diana. (C) Estado de “horquilla” en el que las membranas viral y de la célula diana se encuentran ya fusionadas.

Al formarse el manajo de 6 hélices, las regiones NHR y CHR interaccionan entre sí, con lo que las dos membranas se acercan y se fusionan. Se supone que la energía conformacional liberada tras la formación del manajo de 6 hélices (6-HB) se utiliza para acercar las membranas e inducir la fusión (Melikyan 2008).

Por otro lado, la glicoproteína gp41 que está ocluida en gran medida por la glicoproteína gp120 en el estado nativo del virión, aunque el epítipo de reconocimiento del anticuerpo 2F5 queda accesible (Sattentau et al. 1995), se expone y queda totalmente expuesta al sistema inmunitario en el estado intermedio de pre-horquilla, figura 4.2 (B). La unión de los anticuerpos a gp41 en esta fase impediría la posterior fusión de las membranas, con lo que se bloquearía el proceso de infección en su etapa inicial.

1.4. Estructura de la glicoproteína gp41 y de sus dominios

La glicoproteína gp41 está formada por 345 aminoácidos con una masa molecular de 41 kDa, y no contiene regiones variables definidas. En la secuencia de gp41 se distinguen tres dominios, que están representados en la figura 4.3 (Gabuzda et al. 1992): la región extracelular, también llamada ectodominio, de 172 aminoácidos de longitud; el dominio transmembrana, TM o TMD, formado por 22 residuos; y el tallo citoplasmático, CT, de 151 aminoácidos.

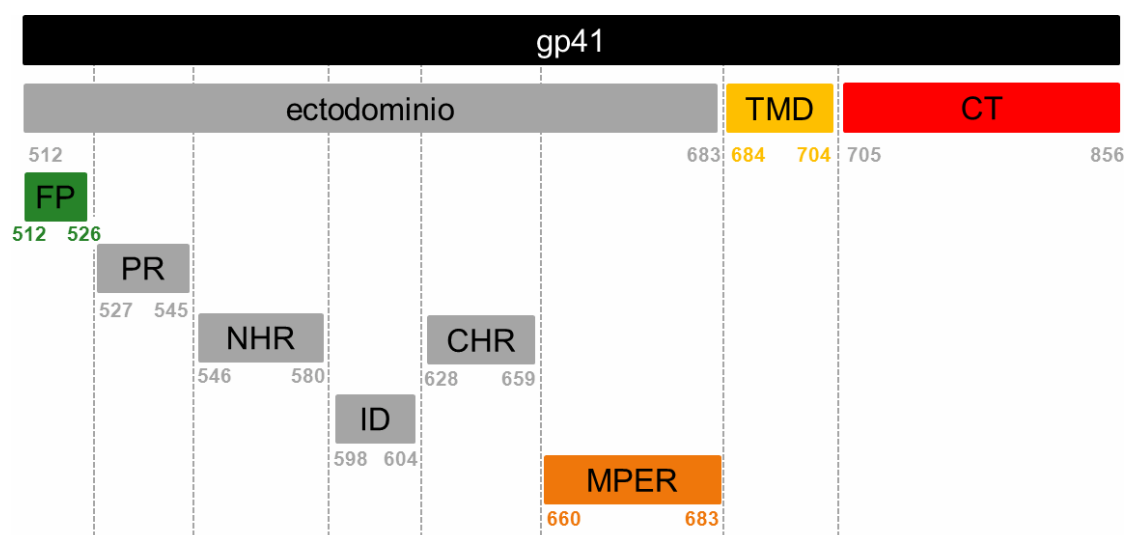


Figura 4.3. Representación esquemática de los dominios, ectodominio, TMD y CT, de la glicoproteína gp41 y de las regiones del ectodominio. Se indican el número de residuo de inicio y final de cada región.

A su vez el ectodominio se encuentra dividido en varias regiones, figura 4.3: (I) FP, la región hidrófoba N-terminal que realiza la función de péptido de fusión; (II) PR, una región polar de 18 residuos de longitud; (III) dos regiones helicoidales denominadas NHR (“N-terminal heptad repeat”), y CHR (“C-terminal heptad repeat”); (IV) el bucle hidrófilo ID (“immunodominant”) que conecta las dos regiones helicoidales, que está formado por siete residuos, y que contiene un enlace disulfuro; y (V) la región MPER (“membrane proximal external región”), que es una región rica en residuos Trp esencial para la infección por VIH-1, y donde se encuentran los epítomos de reconocimiento de varios anticuerpos neutralizantes (sección 1.2).

La estructura tridimensional de gp41 completa no se ha determinado hasta la fecha. Las estructuras publicadas a finales de los 90 corresponden sólo a las regiones NHR y CHR del ectodominio (Chan et al. 1997; Tan et al. 1997; Weissenhorn et al. 1997b). Las estructuras determinadas en estudios más recientes abarcan una mayor parte del ectodominio, pero en ningún caso incluyen la región FP (Buzon et al. 2010; Frey et al. 2010). Todas estas estructuras muestran un manojo de seis hélices (6-HB), que se considera como la estructura del ectodominio una vez se ha producido la fusión, figura 4.2 (C), y que difiere de la estructura nativa meta-estable de gp41 en el clavo viral, así como de la presente en el estado intermedio de “pre-horquilla”.

1.5. La región FP del ectodominio de gp41

La mayoría de los residuos de la región FP, que corresponde a los 15 primeros aminoácidos de gp41, figura 4.4, son Gly, Leu y Phe por lo que es una región muy hidrófoba. Se cree que en el estado no fusogénico la región FP está enterrada en el complejo gp120/gp41, y se expone sólo de forma transitoria después de la unión de gp120 con los receptores CD4. A través de estudios de mutagénesis se ha confirmado que la región FP participa en el proceso por el que se desencadena la fusión de las membranas del virus y la célula (Gallagher 1987; Bosch et al. 1989; Freed et al. 1992; Schaal et al. 1995). En el mecanismo propuesto, la región FP causa la desestabilización de la membrana mediante su inserción oblicua, y a continuación se produce la fusión (Brasseur et al. 1990). En el modelo sugerido para este estado fusogénico (Lorizate et al. 2006b), la región FP se encuentra en contacto con la región MPER como se representa en la figura 4.5.

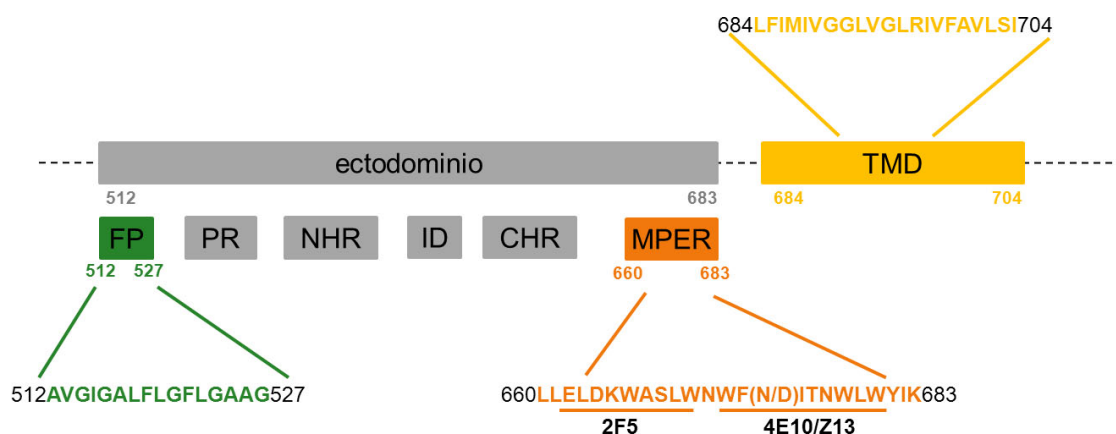


Figura 4.4. Representación esquemática del ectodominio y la región transmembrana (TMD) de la glicoproteína gp41. Se indican las secuencias de las regiones FP (residuos 512–527), en verde; MPER (residuos 660–683), en naranja y TMD (residuos 684–704), en amarillo. Los residuos correspondientes a los epítomos de 2F5 y 4E10/Z13, de la región MPER, están subrayados.

El papel crucial de la región FP en el proceso de fusión está bien establecido, sin embargo, existe controversia acerca de su estructura tanto en la conformación nativa de gp41 como en el estado fusogénico. De hecho, a partir de estudios de FTIR en los péptidos correspondientes a los residuos 512-527 (Martin et al. 1993; Martin & Ruyschaert 1995; Martin et al. 1996) y 512-534 (Nieva et al. 1994; Pereira et al. 1995; Pereira et al. 1997a; Pereira et al. 1997b; Nieva & Agirre 2003; Buzon et al. 2005) se han encontrado estructuras de lámina β , extendidas, helicoidales, desordenadas, agregados β y mezcla de estructuras dependiendo de si las condiciones experimentales empleadas eran en disolución acuosa, DMSO, presencia de iones Ca^{2+} y diferentes modelos de membrana (liposomas, vesículas unilamelares). Mediante RMN en estado sólido (Zheng et al. 2006) se ha visto que la estructura adoptada por la región FP en vesículas puede estar influenciada por la concentración de colesterol. A bajas concentraciones de colesterol, FP adopta una conformación helicoidal, mientras que a concentraciones más altas muestra conformación de hebra. Dado que FP ha demostrado ser fusogénico bajo ambas condiciones, se especula con que la flexibilidad es importante, confiriendo la capacidad de infectar células con membranas compuestas por distintas concentraciones de colesterol y/o para regular el proceso de fusión del VIH-1.

Hasta la fecha se han determinado las estructuras tridimensionales de dos péptidos que contienen la región FP de gp41. La primera publicada es la de un péptido de 37 aminoácidos de longitud que abarca los residuos 512-541 de gp41, y contiene un segmento polar no nativo en su extremo C-terminal, en micelas de SDS (código PDB: 2ARI) (Jaroniec et al. 2005), y la otra para un péptido que corresponde a los residuos 512-535 de gp41, y que también contiene una extensión C-terminal de aminoácidos Lys para aumentar la solubilidad, en

micelas de DPC (códigos PDB: 2PJV) (Li & Tamm 2007). En ambos casos se ha utilizado RMN en disolución, y las estructuras determinadas son de hélice α . Por ello, se ha propuesto que la estructura helicoidal es la conformación activa más probable de la región FP en condiciones fisiológicas.

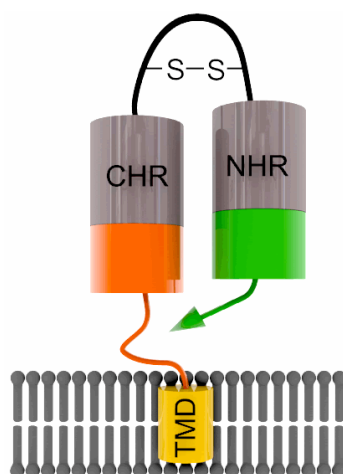


Figura 4.5. Representación de la proteína gp41 unida a la membrana del virus según el modelo de estado fusogénico propuesto por Lorizate et al (Lorizate et al. 2006b). En naranja se representan los residuos 660-683 de la región MPER y en verde los residuos del 512-526, que incluyen la región FP. Las regiones helicoidales, CHR y NHR se indican en color gris.

1.6. La región MPER del ectodominio de gp41

La secuencia de la región MPER de gp41 está estrictamente conservada en la mayoría de los virus VIH-1 aislados hasta la fecha (Muster et al. 1993; Purtscher et al. 1994; Zwick et al. 2001). Esta región juega un papel fundamental en la fusión de la membrana viral con la membrana celular (sección 1.3), como evidencia la completa anulación de la capacidad fusogénica de gp41 si se eliminan segmentos largos de su secuencia (660-670 o 665-682) o la anulación de aproximadamente el 80% si se eliminan segmentos más cortos (666-670, 671-677, 678-682) (Poumbourios et al. 1995; Salzwedel et al. 1999). Además, el péptido que incluye los residuos de 664-683, al asociarse a membranas lipídicas, causa su ruptura (Suarez et al. 2000a; Suarez et al. 2000b). Por otra parte, la región MPER contiene los únicos epítomos de anticuerpos neutralizantes identificados para gp41: 2F5, 4E10 y Z13 (estos dos últimos prácticamente se solapan), figura 4.4, (Muster et al. 1993; Zwick et al. 2001) (apartado 1.2), y es esencial para la actividad inhibitoria de los péptidos derivados de CHR (Jiang et al. 1993b; Jiang et al. 1993a; Wild et al. 1994). Todo esto hace a la región MPER muy atractiva como diana para el diseño de una posible vacuna contra el virus VIH-1.

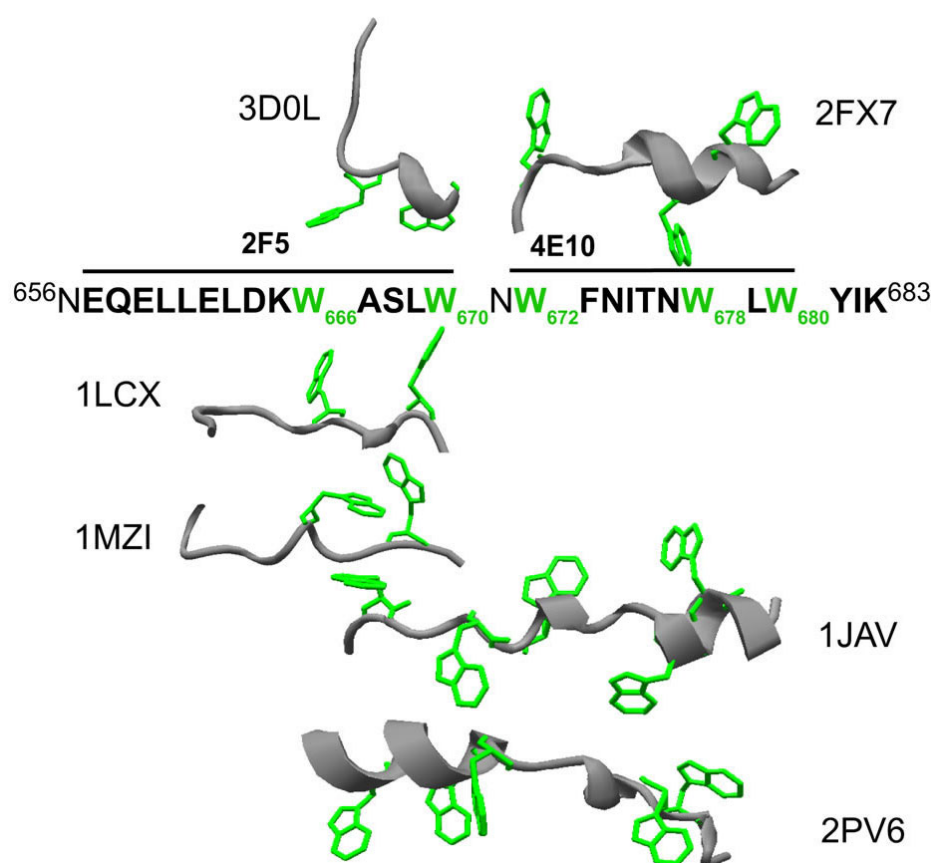


Figura 4.6. Secuencia de la región MPER de gp41 y estructuras de péptidos derivados de MPER en el complejo con los anticuerpos 2F5 y 4E10 (3D0L y 2FX7, en la parte superior); aislados en disolución acuosa (1LCX y 1MZI) y en presencia de micelas de DPC (1JAV y 2PV6). Junto a cada estructura se indica el código de acceso del PDB. Las cadenas laterales de los triptófanos se representan en verde. En la parte superior de la secuencia se indican los epítomos reconocidos por los anticuerpos 2F5 y 4E10 (o Z13).

Como ocurre en el caso de la región FP (apartado 1.5), la implicación y relevancia de la región MPER en el proceso de fusión están establecidas, y, sin embargo, existe controversia acerca de su estructura tanto en la conformación nativa como en el estado fusogénico, ya que en unos casos se describe como una hélice α y en otros como una cadena extendida. Originalmente se predijo que MPER tendría estructura helicoidal (Gallagher et al. 1989), al menos durante el estado fusogénico del virus (Chan et al. 1997; Tan et al. 1997; Weissenhorn et al. 1997b).

Los resultados obtenidos mediante difracción de rayos X y RMN muestran que en disolución acuosa el péptido de 13 aminoácidos correspondiente a los residuos 659-671 tiene una alta población de hélice 3_{10} , 1LCX (Biron et al. 2002), mientras que el correspondiente a los residuos 662-667 solamente presenta conformaciones locales de tipo giro, 1MZI (Barbato et al. 2003), figura 4.6. En presencia de micelas de DPC, se ha determinado mediante RMN que

el péptido que va del residuo 665 al 683 y contiene los 5 triptófanos (Schibli et al. 2001) adopta una hélice en la cual los residuos aromáticos y polares se distribuyen hacia el mismo lado de la hélice, 1JAV, figura 4.6. La estructura del péptido un poco más largo, 662-683, que ha sido determinada por estudios combinados de RMN y EPR también en DPC está formada por dos segmentos unidos a través de una bisagra no helicoidal en $^{673}\text{FN}^{674}$, 2PV6 (Sun et al. 2008), figura 4.6. El primer segmento, residuos 664-672, presenta una conformación de hélice α con dos giros en el extremo C-terminal, mientras que el segundo, residuos 675-683, está menos definido e incluye una hélice extendida 3_{10} .

Por otra parte, se han determinado las estructuras de diversos fragmentos de MPER unidos a los anticuerpos mediante cristalografía de rayos X. Los péptidos correspondientes a los residuos 662-668 (Pai et al. 2000), y 654-670 (Ofek et al. 2004) unidos al anticuerpo 2F5, tienen una conformación extendida con un giro β en los residuos del núcleo del epítipo, $^{664}\text{DKW}^{666}$. El complejo de Fab2F5 unido a un péptido que contiene la región 655-671, 3D0L en la figura 4.6 (Julien et al. 2008), presenta también una conformación bien definida de giro- β en el centro del epítipo enterrado por el parátipo, $^{664}\text{DKW}^{666}$, pero una estructura helicoidal en el extremo C-terminal, que incluye los residuos $^{667}\text{ASLW}^{670}$. La secuencia de reconocimiento del anticuerpo monoclonal MAb4E10, $^{672}\text{WFNITNWLW}^{680}$ en el extremo C-terminal de MPER, unida al anticuerpo Fab4E10 adopta una estructura helicoidal, 2FX7 en la figura 4.6 (Cardoso et al. 2007). Esta región presenta también estructura helicoidal en el complejo de un péptido derivado de MPER (residuos 656-682), con el anticuerpo Fab10E8, que se ha aislado recientemente (Huang et al. 2012).

Esta diversidad estructural de MPER podría estar relacionada con los cambios estructurales que sufre gp41 durante la fusión de membranas. Esto plantea la posibilidad de que la región MPER experimente una transición estructural durante la fusión que vaya desde la forma extendida a la estructura helicoidal (Barbato et al. 2003). Curiosamente, los residuos de MPER que están en contacto directamente con el anticuerpo MAb4E10 (Brunel et al. 2006) son los mismos que en la estructura del péptido 1JAV de MPER interaccionan con las micelas de DPC (Schibli et al. 2001), esto puede deberse a la naturaleza hidrófoba del epítipo 4E10 y apoya que estos residuos interaccionen con las membranas. Por otra parte, se observó que los anticuerpos 4E10 y Z13, que se unen a epítopos superpuestos, figura 4.4, interaccionan con dos caras diferentes de la hélice de MPER (Nelson et al. 2007). Todas estas observaciones apoyan la naturaleza flexible de la región MPER y sugieren la existencia de más de una estructura relevante en la neutralización del virus por los anticuerpos. Esta variabilidad estructural implica que inmunógenos distintos con diferentes orientaciones de la

hélice MPER, o incluso con estructuras extendidas, puedan ser útiles para generar anticuerpos neutralizantes.

1.7. La región TMD de gp41

La secuencia de la región TM o TMD, que ancla el complejo Env a la bicapa lipídica, consta prácticamente en su totalidad de residuos hidrófobos, figura 4.4. Esta región está altamente conservada entre las diferentes cepas aisladas del VIH-1. Dado que la sustitución de esta región por otros anclajes lipídicos impide la fusión de las membranas (Salzwedel et al. 1993; Weiss & White 1993; Lin et al. 2003; Miyauchi et al. 2005), se ha propuesto que desempeña un papel fundamental en la fusión de las membranas viral y celular (Welman et al. 2007). A pesar de su relevancia, probablemente como consecuencia de los problemas asociados con la expresión y purificación de proteínas con regiones transmembrana, no se ha resuelto hasta la fecha la estructura de ninguna construcción de gp41 que incluya la región TMD, ni de la región TMD por sí sola. Los modelos realizados computacionalmente sugieren que adopta forma helicoidal (Gallagher et al. 1989).

1.8. Objetivo

La infección por el virus VIH-1 conlleva una etapa inicial de fusión de las membranas del virus VIH-1 y de la célula hospedadora, cuyo mecanismo detallado se desconoce. Es, sin embargo, bien conocido que en este proceso de fusión interviene el complejo Env presente en la membrana viral VIH-1, figura 4.2, y que está formado por las glicoproteínas gp120 y gp41. Varios anticuerpos monoclonales neutralizantes de VIH-1 reconocen epítomos localizados en estas proteínas, que son lineales en el caso de gp41 y discontinuos en gp120. Se ha propuesto que los péptidos que contienen uno o más de los epítomos podrían ser utilizados como vacunas efectivas frente a VIH. Para este fin, los epítomos de gp41 son más adecuados que los de gp120, dada la conservación de su secuencia, así como por la estructura lineal y continua de los mismos.

En este contexto, el objetivo general planteado en este capítulo es contribuir al entendimiento de: (I) el proceso de fusión de las membranas virus-célula hospedadora, y (II) la relación entre la estructura y la capacidad inmunogénica de péptidos derivados del ectodominio de la glicoproteína gp41.

Para ello se ha procedido al diseño de varios péptidos derivados de gp41 y a la determinación de sus estructuras mediante RMN en disolución. Dada la naturaleza hidrófoba de los péptidos y que su entorno natural es la región próxima a la membrana o la propia membrana, se han utilizado dos condiciones experimentales diferentes que emulan los entornos de membrana, y que se utilizan frecuentemente para los estudios de RMN en disolución. Por una parte, se ha empleado un medio acuoso con cierto carácter apolar, en concreto, una mezcla hexafluoroisopropanol/agua 25:75 v/v (25% HFIP), y por otra un medio micelar, en particular, micelas de dodecilsulfocolina (20 mM DPC), ya que la fosfatidilcolina es uno de los principales constituyentes de la bicapa lipídica de las membranas celulares.

Los resultados obtenidos para el péptido MPERp ya han conducido a la publicación de un artículo en colaboración con el grupo del Prof. J. L. Nieva, Departamento de Bioquímica y Biología Molecular, Universidad del País Vasco (UPV/EHU): Serrano S., Araujo A., Apellániz B., Bryson S., Carravilla P., De la Arada I., Huarte N., Rujas E., Pai E.F., Arrondo J.L., Domene C., Jiménez M.A., Nieva J.L. *“Structure and immunogenicity of a peptide vaccine including the complete HIV-1 gp41 2F5 epitope. Implications for antibody recognition mechanism and immunogen design”* J. Biol. Chem. **(2014)**. Dos artículos relativos a los demás péptidos se encuentran en fase de redacción.

2. MATERIALES Y MÉTODOS

2.1. Preparación de las muestras

Los péptidos estudiados fueron producidos a través de la técnica de síntesis en fase sólida y purificados por HPLC de fase inversa por el grupo del Dr. David Andreu (Universidad Pompeu Fabra, Barcelona), en el caso de los péptidos HybK3 and ScrHybK3, y por el grupo del Dr. José L. Nieva (Unidad de Biofísica, Departamento de Bioquímica de la Universidad del País Vasco), en el caso de los péptidos MPERp, CpreTM y TMDp. Los péptidos correspondientes a las regiones MPER y TMD se denominan como MPERp y TMDp para evitar confusiones entre los péptidos y los dominios de gp41.

Para cada péptido se prepararon dos muestras: una en un disolvente mixto alcohol fluorado/agua, en concreto 25% 1,1,1,3,3,3,3-hexafluoro-isopropanol, HFIP-d2 deuterado al 98%, (Cambridge Isotope Laboratories, DLM-143-5X1), y la otra en un medio micelar, 20 mM dodecilsulfocolina, DPC-d38, (Cambridge Isotope Laboratories, DLM-2274-0.5). Para preparar la muestra 25% HFIP, se disolvió aproximadamente 1 mg de péptido liofilizado en 0.38 mL de una solución acuosa H₂O/D₂O 9:1 v/v que contenía tampón HEPES 2 mM a pH 7.0, y se añadieron 0.13 mL de HFIP. Las muestras en micelas de DPC se han preparado disolviendo aproximadamente 1 mg del péptido liofilizado en 0.5 mL de una disolución 20 mM DPC deuterado y 2 mM HEPES pH 7.0 en H₂O/D₂O 9:1 v/v. El volumen final de cada muestra fue de 0.5 mL y la concentración de péptido de aproximadamente 0.5-0.7 mM. De aquí en adelante las muestras al 25% de 1,1,1,3,3,3,3-hexafluoro-isopropanol y en 20 mM dodecilsulfocolina se nombrarán como 25% HFIP y 20 mM DPC, respectivamente.

El pH se midió con un microelectrodo de vidrio Hamilton de 3 mm de diámetro, y se ajustó cuando fue necesario a pH 7.0 mediante la adición de alícuotas de DCl o NaOD.

Todas las muestras llevan DSS (2,2-dimetil-2-silapentano-5-sulfonato sódico) como referencia interna de los desplazamientos químicos de ¹H a una concentración aproximada de 0.1 mM.

2.2. Adquisición de los espectros de RMN

Para la adquisición de los espectros de RMN se utilizó un espectrómetro de RMN Bruker Avance 600, con un campo magnético de 14.1 Teslas que opera a 600.13 MHz para el protón, equipado con una criosonda TXI-“Criosonda Triple Inversa”-(^1H , ^{13}C , ^{15}N).

Para todas las muestras se han adquirido los experimentos bidimensionales homonucleares, ^1H - ^1H COSY (Aue et al. 1976), ^1H - ^1H TOCSY (Braunschweiler & Ernst 1983) y ^1H - ^1H NOESY (Jeener et al. 1979; Kumar et al. 1980; Macura & Ernst 1980), utilizando secuencias estándar de pulsos y se ha realizado la eliminación de la señal del agua mediante presaturación. El tiempo de mezcla empleado para los experimentos ^1H - ^1H NOESY fue de 150 ms y para los espectros ^1H - ^1H TOCSY de 60 ms.

También se realizaron experimentos de correlación heteronuclear ^1H - ^{13}C HSQC, con ^{13}C en abundancia natural.

Los espectros se han adquirido a 25 °C y la temperatura de la muestra dentro de la sonda de RMN se calibró utilizando una muestra de metanol (Duerst & Verbach 1965).

Para comprobar la estabilidad de los péptidos durante la adquisición de los experimentos de RMN se comprobó que no había cambios entre los espectros monodimensionales de ^1H adquiridos antes y después de cada espectro bidimensional.

2.3. Procesado de los espectros de RMN

Los espectros de RMN se procesaron con el programa Topspin 3.0 (Bruker Biospin, Karlsruhe, Alemania). En el caso de los bidimensionales, tras la aplicación de la transformada de Fourier se realizó la corrección de línea base en la dimensión directa e indirecta según fuera necesario.

2.4. Asignación espectral

Para la visualización y análisis de los espectros bidimensionales de RMN se utilizaron los programas SPARKY (Goddard 2005) y CCPNMR Analysis (Vranken et al. 2005).

La asignación espectral se ha realizado siguiendo el protocolo que se describe en la sección 2.6 del capítulo III.

2.5. Determinación de la estructura secundaria

De manera similar a lo apuntado en el caso de la asignación espectral, la determinación de los elementos de estructura secundaria de los péptidos se ha realizado como se describe en el apartado 2.7 del capítulo III.

2.6. Cálculo de estructuras tridimensionales

La determinación de la estructura tridimensional de una biomolécula consiste en la obtención de las coordenadas atómicas compatibles con las restricciones geométricas experimentales procedentes de los parámetros de RMN (restricciones de distancia, angulares, y de orientación).

Las estructuras de los péptidos derivados de gp41 estudiados en este capítulo se han determinado utilizando las restricciones de distancia obtenidas a partir de las señales NOE en el espectro ^1H - ^1H NOESY, y de las restricciones de ángulos diedros Φ y ψ obtenidas a partir de los desplazamientos químicos de $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ usando el programa TALOS (Cornilescu et al. 1999).

El cálculo se ha realizado utilizando el programa CYANA 2.1 (Combined assignment and dYnamics Algorithm for NMR Applications) (Güntert 2004) y siguiendo un protocolo que consta de tres etapas. En la primera se aplica el procedimiento iterativo de asignación automática de correlaciones NOE con el programa CYANA 2.1, que consta de siete ciclos en cada uno de los cuales se calculan 100 confórmeros. La lista de restricciones de distancia resultante del último ciclo se revisó y comprobó mediante el análisis del espectro NOESY, y las restricciones ambiguas se eliminaron o relajaron para generar la lista definitiva de restricciones de distancia. La segunda etapa consiste en el cálculo de 100 confórmeros utilizando el procedimiento estándar de “simulated annealing” de CYANA 2.1 y como restricciones experimentales la lista definitiva de restricciones de distancia obtenida en la etapa anterior y las restricciones de ángulos diedros Φ y ψ . El conjunto de los 20 confórmeros con los valores de función blanco más bajas se sometieron finalmente a una minimización de energía que en el caso de los péptidos HybK3 y ScrHybK3 se realizó dentro del programa CYANA 2.1, y en los péptidos MPERp, CpreTM y TMD utilizando el modelo generalizado de Born de solvatación continua implementado en el programa AMBER9 (Case 2006).

2.7. Visualización y análisis de las estructuras

Las estructuras calculadas para los péptidos han sido examinadas y representadas utilizando el programa MOLMOL (Koradi et al. 1996), PROCHECK/NMR (Laskowski et al. 1996) y el servidor de validación de estructura de proteínas PSVS server(<http://psvs-1.4-dev.nesg.org/>).

3. RESULTADOS

3.1. Péptidos quimera de las regiones FP y MPER: HybK3 y scrHybK3

3.1.1. Diseño de los péptidos

Con el fin de encontrar evidencias experimentales que apoyen la hipótesis de que las regiones FP y MPER del ectodominio de gp41 interaccionan entre sí en la proximidad de la membrana durante la fusión de las membranas del virus y de la célula hospedadora (de la Arada et al. 2009) (ver figura 4.5 y apartados 1.3 y 1.5), se procedió al diseño de dos péptidos quimera o híbridos, denominados HybK3 y scrHybK3. El péptido HybK3, representado en la figura 4.7, contiene las secuencias de la región FP (⁵¹⁴GIGALFLGFLGAAGS⁵²⁸) y de la región de MPER en la que se localiza el epítipo de reconocimiento del anticuerpo monoclonal neutralizante 2F5 (⁶⁵⁵KNEQELLELDKWASLWN⁶⁷¹) unidas mediante un separador flexible que sustituye a los 127 residuos que separan las regiones FP y MPER en la proteína gp41 completa (Lorizate et al. 2006b). El separador flexible está formado por dos residuos Lys y el ácido 6-amino hexanoico, Ahx. Los residuos de la región FP se han seleccionado teniendo en cuenta que los péptidos sintéticos cortos derivados del dominio N-terminal de la región FP de gp41 son capaces de promover la mezcla de lípidos y la fusión de las membranas lipídicas (Durell et al. 1997; Epand 2003; Nieva & Agirre 2003). El péptido scrHybK3 tiene los mismos residuos que HybK3, pero se ha modificado el orden ("scrambled") de los correspondientes a la secuencia central de FP, donde se encuentran la mayor parte de los residuos hidrófobos Leu y Phe, figura 4.7. Es de esperar que la alteración en la secuencia de la región FP afecte de forma significativa la posible interacción específica entre la región FP y la región N-terminal de MPER.

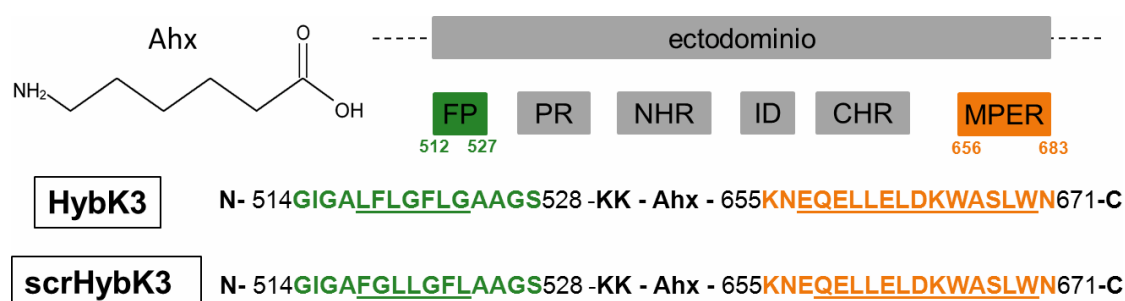


Figura 4.7. Secuencias de los péptidos HybK3 y scrHybK3. Los residuos derivados de la región FP se indican en verde, donde los residuos de la región central hidrófoba de FP se han subrayado, y los derivados de la región MPER en naranja, cuyos residuos subrayados corresponden al epítipo de 2F5. En la parte superior a la derecha se representan de forma esquemática todos los dominios del ectodominio de la proteína gp41, y a la izquierda la molécula de ácido 6-amino hexanoico (Ahx).

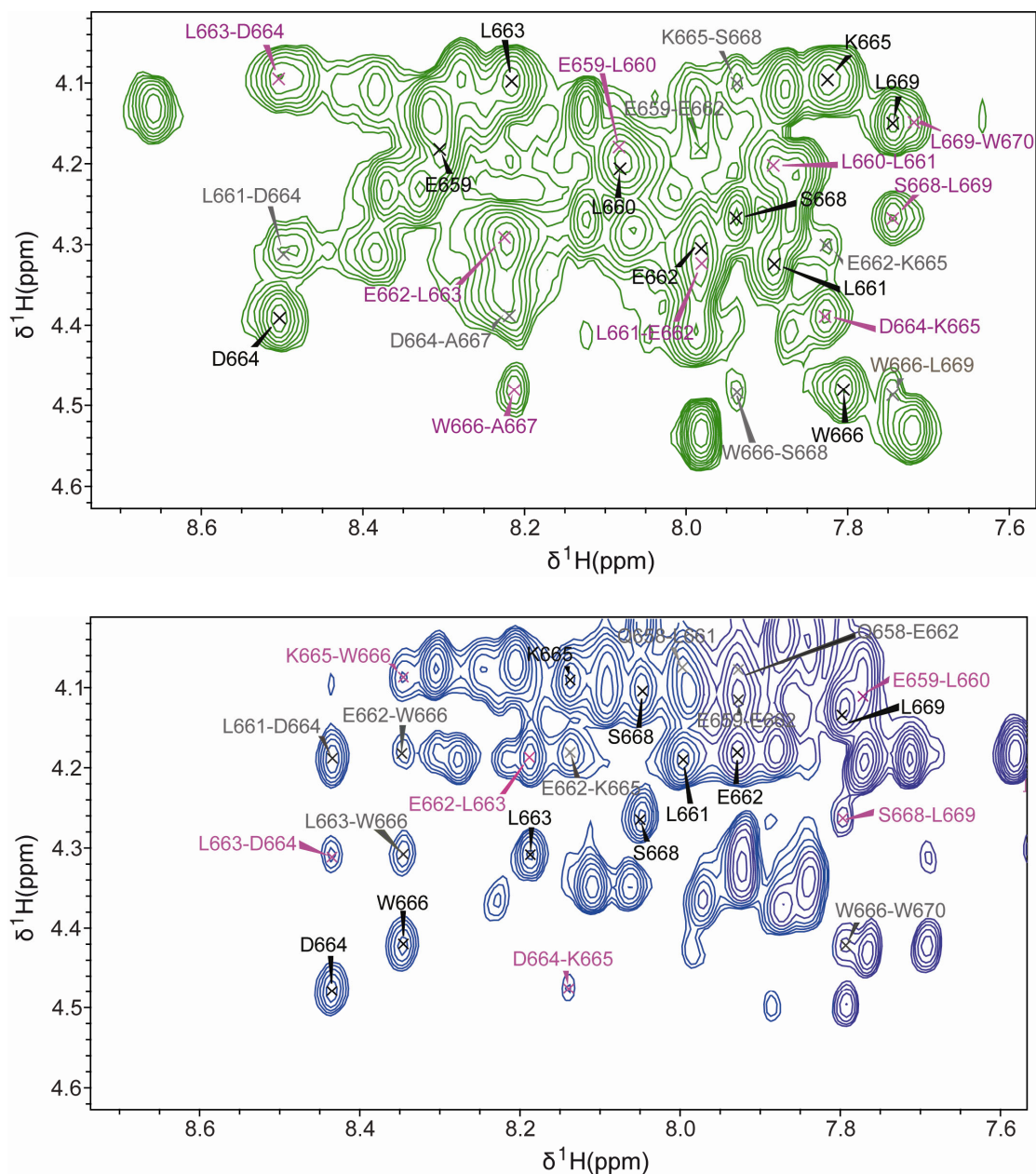


Figura 4.8. Región del experimento ^1H - ^1H NOESY del péptido HybK3 en 20 mM DPC, en la parte superior, y en 25% HFIP, en la parte inferior, a pH 6.8 y 25 °C. Las etiquetas en negro muestran los picos de cruce correspondientes a las correlaciones intrarresiduales, en magenta las secuenciales y en gris las no secuenciales de los residuos que van de Leu661 a Trp669.

3.1.2. Espectros de RMN

Los espectros de RMN de los péptidos HybK3 y scrHybK3 se han adquirido en 25% HFIP y en 20 mM DPC (ver apartado 2.1) a 25 °C. Estos medios se aproximan a las condiciones del estado de “horquilla” previo a la fusión de las membranas (apartados 1.3 y 1.5; y figuras 4.2 C y 4.4). La buena calidad de los espectros adquiridos en 25% HFIP, tanto para el péptido HybK3 como para scrHybK3, ha permitido la asignación completa de los protones, y de los carbonos (tablas A.1 y A.2 del apéndice). Sin embargo, el medio micelar (20 mM DPC) conduce a un ligero ensanchamiento de las señales de resonancia en los dos péptidos, lo que disminuye la relación señal/ruido de los espectros de RMN, como se observa en la parte superior de la figura 4.8. Por ello, ha sido posible la asignación prácticamente completa de los protones (tablas A.3 y A.4 del apéndice) mientras que los carbonos sólo se han podido asignar parcialmente. Además del ensanchamiento de las señales, hay que tener en cuenta que los espectros ^1H - ^{13}C -HSQC se han adquirido en abundancia natural del ^{13}C lo que limita la sensibilidad del experimento.

3.1.3. Análisis cualitativo de la estructura secundaria

Una vez asignados los espectros de RMN de los péptidos HybK3 y scrHybK3, los desplazamientos conformacionales de los $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ (apartado 2.5) permiten identificar la estructura secundaria adoptada por los péptidos. Se excluyen de este análisis los aminoácidos N- y C-terminales. Los desplazamientos conformacionales de $^1\text{H}_\alpha$ ($\Delta\delta_{\text{H}\alpha}$) observados para los dos péptidos tanto en presencia de HFIP como en el medio micelar, figura 4.9, presentan valores negativos y grandes en valor absoluto ($\Delta\delta \geq 0.05$ ppm) tanto para la región derivada del dominio FP (residuos 516-528), como para la del dominio MPER (residuos 657-669), que son indicativos de que ambas regiones adoptan estructura de hélice α . En presencia de las micelas de DPC, la hélice de la región FP se extiende a los dos residuos no nativos de Lys. Algunos residuos dentro de estas regiones presentan valores pequeños en magnitud, que podrían indicar una cierta irregularidad en la hélice α formada, o deberse a efectos de anisotropía debidos a los anillos aromáticos. Esto ocurre en los dos péptidos en la región derivada del dominio MPER para el aminoácido Leu663 en 25% HFIP, y para los residuos Leu661 y Glu662 en las micelas de DPC. En el caso de la región derivada del dominio FP, la única excepción a esta tendencia conformacional en el péptido HybK3 es el residuo Gly521. En la región N-terminal, el péptido scrHybK3 presenta excepciones para el aminoácido Leu520 en 25% HFIP, y para Gly519 en presencia de las micelas de DPC.

En presencia de HFIP, la formación de estructuras helicoidales en estas dos regiones de ambos péptidos viene corroborada por los desplazamientos conformacionales de los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$. Los valores de $\Delta\delta_{\text{C}_\alpha}$ son positivos y grandes en magnitud ($(\Delta\delta \geq 0.5 \text{ ppm})$), y los valores de $\Delta\delta_{\text{C}_\beta}$ son negativos y grandes en valor absoluto (figura A.1 del apéndice).

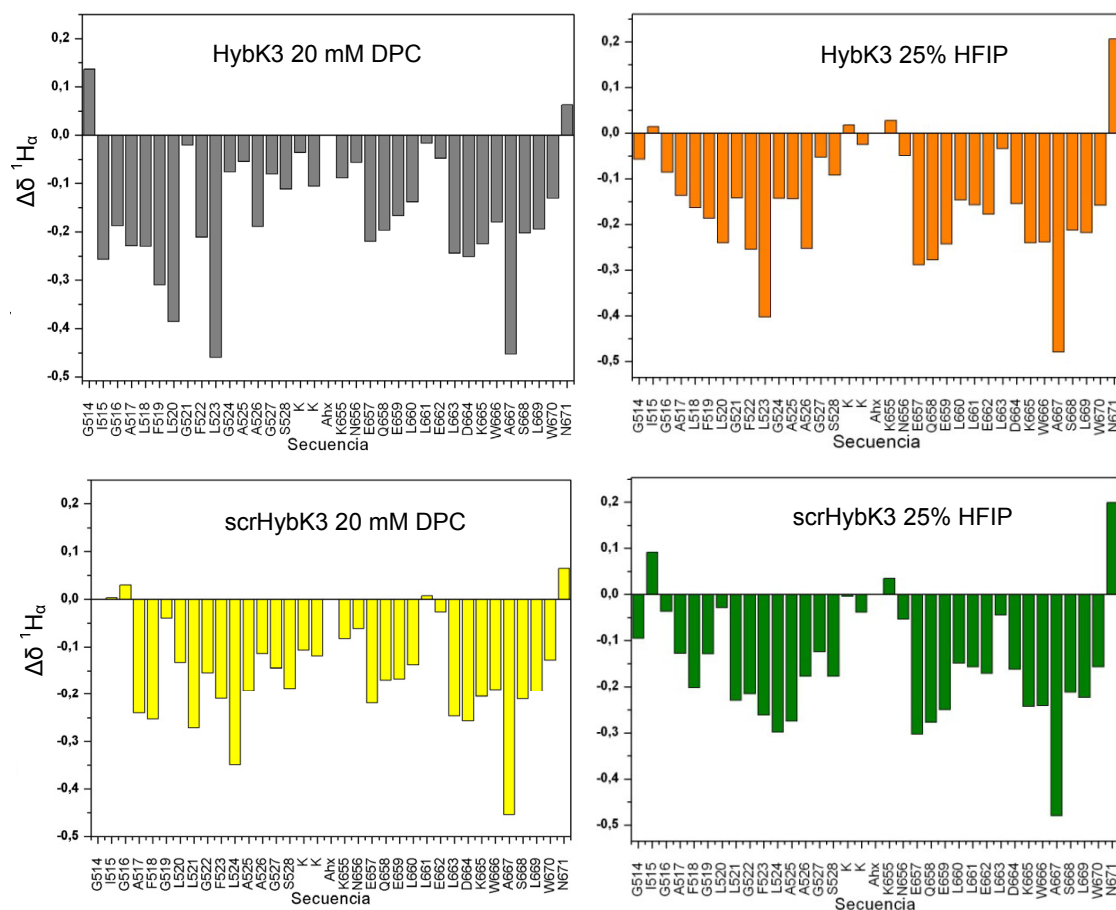


Figura 4.9. Desplazamientos conformacionales de $^1\text{H}_\alpha$ ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm) para el péptido HybK3, en la parte superior, y scrHybK3 en la parte inferior a pH 6.8 y 25 °C en 20 mM DPC a la izquierda y en 25% HFIP a la derecha.

La formación de hélice α en las regiones FP y MPER de los dos péptidos se confirma por el conjunto de señales de correlación NOE observadas, figura 4.10, en particular, las no secuenciales de tipo $d_{\alpha\text{N}(i,i+2)}$, $d_{\alpha\text{N}(i,i+3)}$, $d_{\alpha\text{N}(i,i+4)}$, $d_{\alpha\beta(i,i+3)}$ y $d_{\text{NN}(i,i+2)}$, todas ellas características de hélice α (tabla 3.1, capítulo III). Algunas de estas correlaciones NOE se pueden ver en las regiones de los espectros ^1H - ^1H NOESY mostrados en la figura 4.8.

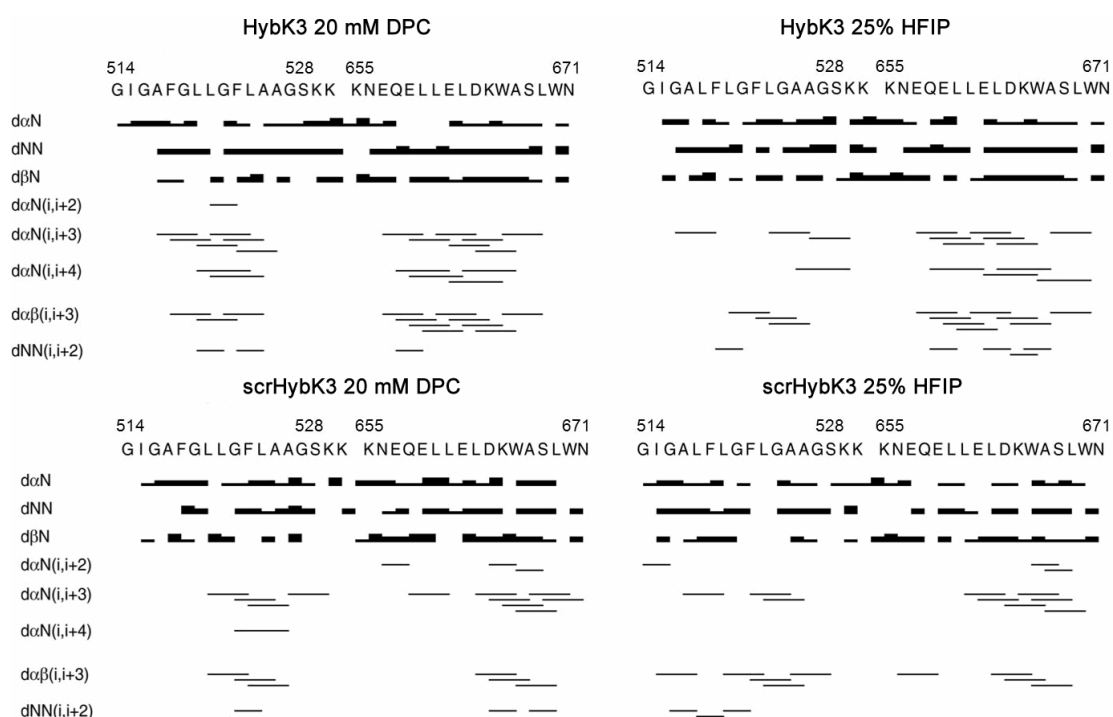


Figura 4.10. Resumen de los NOEs observados para los péptidos HybK3 y scrHybK3.

3.1.4. Estructuras tridimensionales de los péptidos

Dado que la magnitud de los desplazamientos conformacionales, figura 4.9, es indicativa de poblaciones elevadas de estructura, con el fin de visualizar y poder analizar en detalle las características de las conformaciones adoptadas por los péptidos HybK3 y scrHybK3 tanto en 25% HFIP como en 20 mM DPC se han calculado las estructuras siguiendo el protocolo de cálculo descrito en el apartado 2.6. El número de restricciones de distancia y de restricciones angulares utilizado en cada caso se indica en la tabla 4.1. Las estructuras resultantes cumplen todas las restricciones experimentales, ya que no presentan violaciones ni angulares ni de distancia, y su calidad es buena, como indica el hecho de que ni la longitud ni los ángulos de enlace se desvían de las geometrías ideales, y que los ángulos Φ y ψ se encuentran en las zonas favorables en el mapa de Ramachandran.

		Péptido Hybk3		Péptido scrHybk3	
		DPC	HFIP	DPC	HFIP
Restricciones de distancia	Rango Corto ($i - j \leq 1$)	296	210	226	210
	Rango Medio ($1 < i - j < 5$)	119	68	67	68
	Rango Largo ($ i - j \geq 5$)	0	0	0	0
	Número Total	415	278	293	278
	Restricciones por residuo	12.58	8.42	8.88	8.42
Número de restricciones angulares	Ángulos ϕ	31	32	25	32
	Ángulos ψ	29	30	25	30
	Número Total	60	62	50	32
Violaciones máximas por estructura	Distancia (Å)	0.0	0.0	0.0	0.0
	Ángulos Diedros ($^{\circ}$)	0.0	0.0	0.0	0.0
Valores de energía de las estructuras	Función blanco de CYANA	1.39±	1.39±	1.44±	1.39±
		6.86E-04	2.47E-03	3.53E-03	2.47E-03
Desviaciones de la geometría ideal	Longitud de enlaces (Å)	0.0002	0.0009	0.0024	0.0009
	Ángulos de enlace ($^{\circ}$)	0.017	0.041	0.431	0.041
Mapa de Ramachandran (%)	Residuos en regiones más favorables	96.5	95.0	81.0	95.0
	Residuos en regiones permitidas adicionales	3.5	5.0	17.1	5.0
	Residuos en regiones permitidas generosamente	0.0	0.0	1.7	0.0
	Residuos en regiones no permitidas	0.0	0.0	0.2	0.0

Tabla 4.1. Datos estadísticos de las 20 estructuras de menor energía para los péptidos HybK3 y scrHybK3 en 20 mM DPC y 25% HFIP, respectivamente.

Considerando todos los residuos, ninguna de las cuatro familias de estructuras calculadas se encuentra bien definida, como se puede apreciar para el péptido HybK3 en la figura 4.11 (A) y para el péptido scrHybK3 en 4.12 (A), y como indican los valores elevados de RMSD global, tabla 4.2. Sin embargo, los residuos de la región FP por sí solos se encuentran bien definidos, figuras 4.11 (B) y 4.12 (B), e igualmente, e incluso en mayor medida, lo están los residuos de la región MPER por sí solos, figuras 4.11 (C) y 4.12 (C). En todos los casos las dos regiones adoptan estructuras helicoidales que parecen ser independientes y no interaccionar entre sí en ninguno de los dos péptidos.

La media de los ángulos entre las hélices FP y MPER para el péptido HybK3 en DPC es de 104 ± 41 y 96 ± 49 en HFIP, mientras que en el péptido scrHybK3 en DPC es de 80 ± 47 y en HFIP 97 ± 39 . El valor medio de ambos péptidos en HFIP es bastante similar y varía más en presencia de DPC, pero estas similitudes y diferencias no son significativas dado que en todos los casos los ángulos entre las hélices presentan un rango muy amplio de variación, de

11 a 170°. Es decir, en unas estructuras las dos hélices se sitúan en posiciones prácticamente antiparalelas y en otras casi paralelas. Esta variabilidad de los ángulos entre las hélices es el resultado de que no se observan NOEs entre ambas regiones, y es la responsable de los elevados valores de RMSD que muestra la estructura global, tabla 4.2.

HybK3		Global (514-671)	Hélice FP (514-530)	Hélice MPER (655-671)
20 mM DPC	Esqueleto	6.2 ± 2.5	1.4 ± 0.6	0.4 ± 0.3
	Átomos pesados	7.4 ± 2.6	2.1 ± 0.6	1.6 ± 0.4
25% de HFIP	Esqueleto	6.6 ± 2.4	1.5 ± 0.4	0.7 ± 0.2
	Átomos pesados	7.7 ± 2.4	2.1 ± 0.4	1.7 ± 0.3
scrHybK3				
20 mM DPC	Esqueleto	6.7 ± 1.7	3.6 ± 0.9	1.3 ± 0.7
	Átomos pesados	7.6 ± 1.7	4.7 ± 1.0	2.0 ± 0.7
25% de HFIP	Esqueleto	5.6 ± 2.0	2.2 ± 0.6	0.5 ± 0.2
	Átomos pesados	6.8 ± 1.9	2.8 ± 0.6	1.6 ± 0.3

Tabla 4.2. Valores de RMSD en Å para las regiones FP y MPER de los péptidos HybK3 y scrHybK3 en 20mM DPC y 25% HFIP. Valores obtenidos con MOLMOL (Koradi et al. 1996).

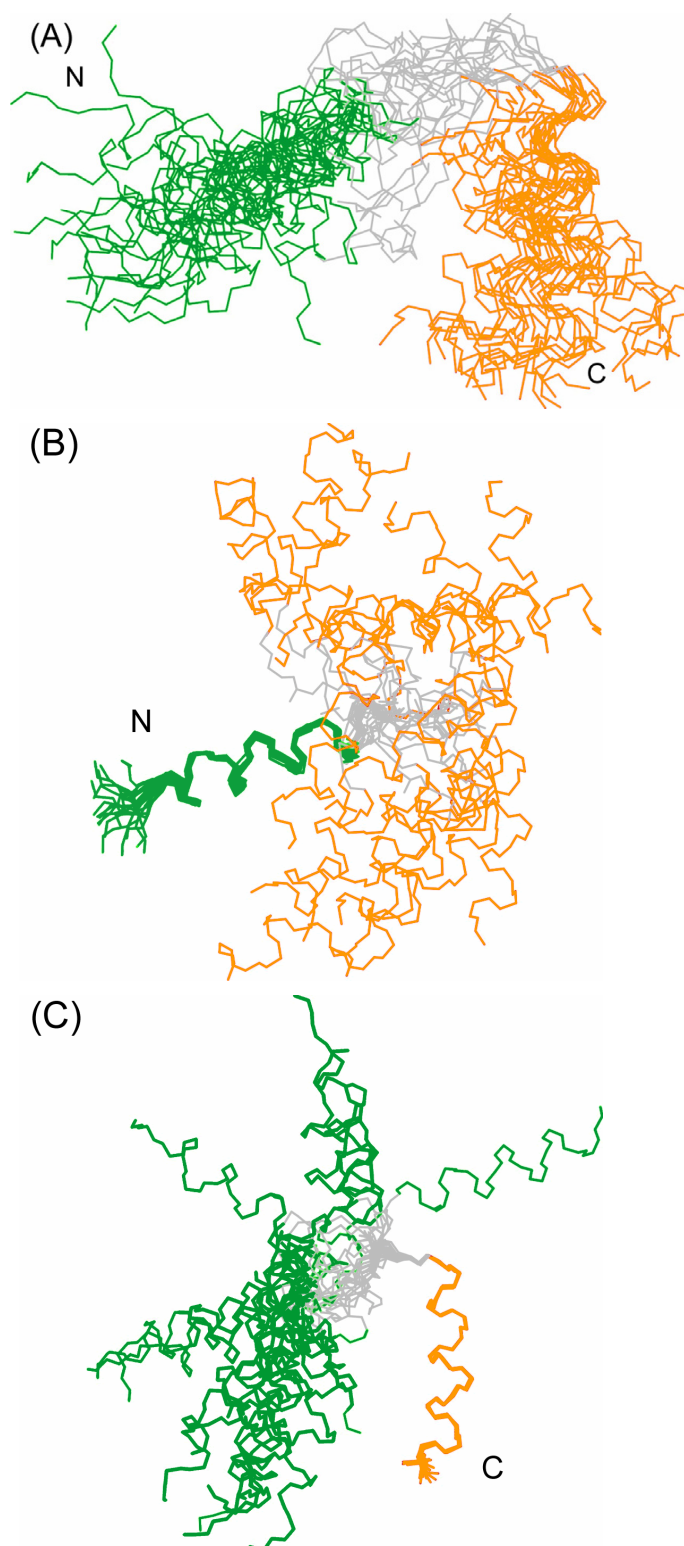


Figura 4.11. (A) Superposición de las 20 estructuras calculadas para HybK3 en 25% HFIP, en gris el separador flexible Ahx, en verde FP y en naranja MPER. (B) Ajuste de la región FP del extremo N-terminal y (C) de la región MPER del extremo C-terminal.

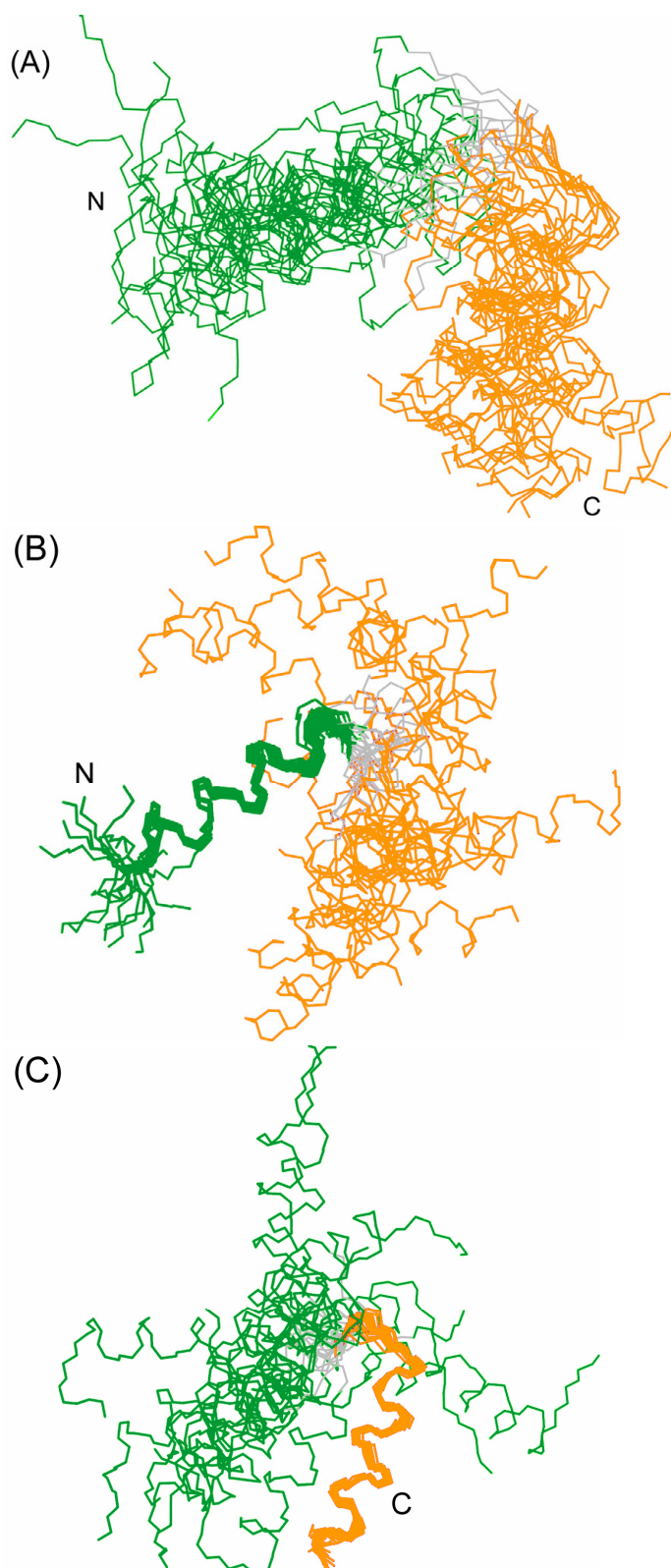


Figura 4.12. (A) Superposición de las 20 estructuras calculadas para srcHybK3 en 25% HFIP. (B) Ajuste de la región FP del extremo N-terminal y (C) Ajuste de la región MPER del extremo C-terminal.

En el péptido HybK3 la región FP tiene estructura de hélice tanto en medio micelar como en HFIP, figura 4.13, lo que varía de un medio a otro es la longitud de estas hélices. La región FP en DPC tiene una estructura helicoidal que va de Ile515 hasta Ala525, donde los residuos 515-522 muestran estructura de hélice α , y los residuos 523-525 de hélice 3_{10} . Por tanto en los residuos $^{522}\text{FL}^{523}$ se produce una variación de la estructura secundaria, y un cambio en la dirección del eje de la hélice. En la región FP en HFIP, la estructura helicoidal es más regular, además de ligeramente más larga ya que comprende los residuos que van de Ala517 a Lys530.

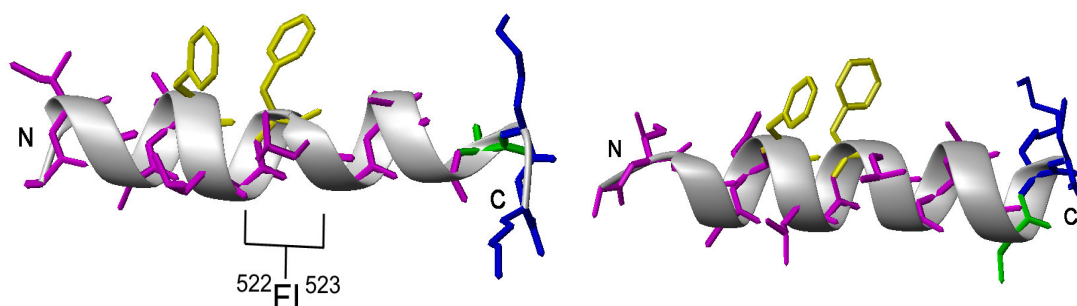


Figura 4.13. Estructura de la región FP del péptido HybK3. La hélice en medio micelar a la izquierda y en 25% HFIP a la derecha. Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu, Gly y Ala (residuos no polares); en verde a Ser (residuos polares neutros); en azul a Lys (residuo polar básico); y en amarillo a Phe (residuos aromáticos).

En el caso de la región MPER, que prácticamente corresponde con el epítipo de 2F5, del péptido HybK3 en DPC la hélice va de Glu657 a Ala667 y en HFIP de Glu657 a Leu669. En ambos casos la estructura helicoidal sufre un cambio en la dirección del eje de la estructura en los residuos $^{666}\text{WA}^{667}$, figura 4.14, que corresponde con el paso de hélice α a hélice 3_{10} . Resulta de interés la observación de que cuando el péptido HybK3 está en micelas de DPC los residuos aromáticos, Trp666 y Trp670 quedan mirando hacia la misma dirección.

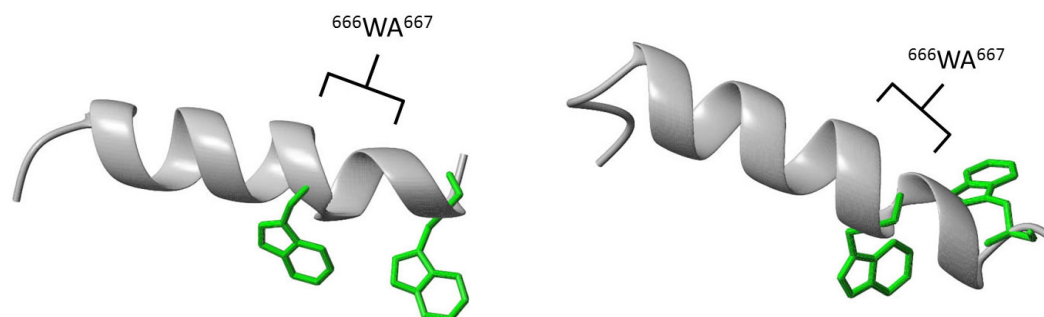


Figura 4.14. Hélice α de la región MPER de HybK3 que va de Glu657 a Leu669, a la izquierda en 20 mM DPC y a la derecha en 25% HFIP. En ambos medios se produce un giro en la transición de hélice α a hélice 3_{10} en los residuos $^{666}\text{WA}^{667}$. Las cadenas laterales de los residuos de Trp666 y Trp670 están representadas en verde.

En el péptido scrHybK3 la región scrFP muestra estructura helicoidal pero cuando se estudia en presencia de micelas de DPC esta hélice sólo incluye los residuos que van de Leu521 hasta Ala526, por lo que sufre un acortamiento significativo con respecto a la hélice que se forma en presencia de HFIP (que va de Ala517 hasta Lys530) y con respecto a las hélices determinadas en el péptido HybK3, figura 4.15.

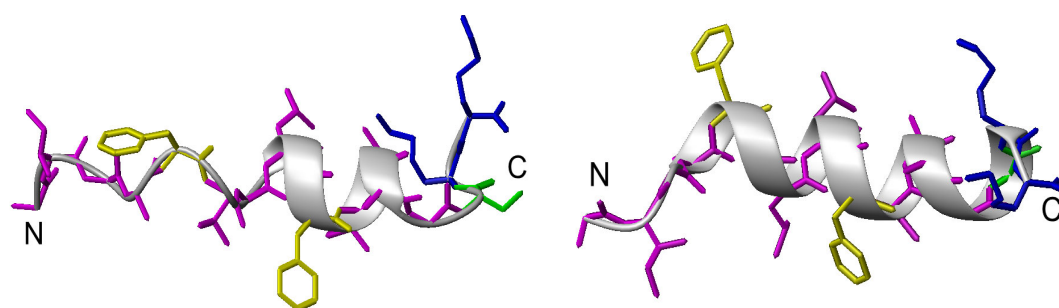


Figura 4.15. Estructura de la región FP en el péptido scrHybK3 en presencia de 20 mM DPC a la izquierda y en 25% HFIP a la derecha. Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu, Gly y Ala; en verde oscuro a Ser; en verde claro con Lys; y en amarillo a Phe.

La región MPER del péptido scrHybK3 muestra una estructura helicoidal con variaciones mínimas ya que en DPC va de Glu657 a Trp670 y en HFIP de Glu657 a Ser668, figura 4.16. En el caso de la estructura en DPC, como ocurre en HybK3, la hélice sufre una desviación en su eje en los residuos ⁶⁶⁵KW⁶⁶⁶. Este giro se encuentra en el cambio de hélice α , que va de Glu657 a Lys665, a la hélice 3_{10} , que va del Trp666 a Trp670.

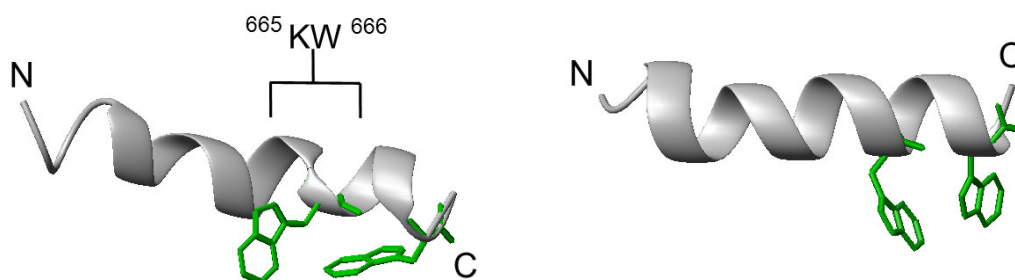


Figura 4.16. Estructura de la región MPER del péptido scrHybK3 en presencia de micelas de DPC a la izquierda y en 25% HFIP a la derecha. En DPC se produce un giro en los residuos ⁶⁶⁵KW⁶⁶⁶ en el paso de hélice α a hélice 3_{10} . Las cadenas laterales de los triptófanos están representadas en verde.

Tanto en el péptido HybK3 como en scrHybK3, las hélices de las regiones FP y MPER son más largas en 25% HFIP que en el medio con micelas de DPC.

3.2. Péptidos de la regiones MPER y TMD

3.2.1. Diseño de los péptidos MPERp, CpreTM y TMDp

Dada la relevancia biológica del dominio MPER de gp41 tanto en la fusión de membranas como en el reconocimiento por anticuerpos, existen numerosos estudios estructurales de fragmentos de esta región (apartado 1.6). En el presente estudio se siguió un criterio basado en los estudios de inmunogenicidad y se diseñaron los péptidos MPERp y CpreTM, figura 4.17, ambos de 28 residuos de longitud. El péptido MPERp, que corresponde a los residuos 656-683 de gp41, incorpora todo el extremo N-terminal que contiene el epítipo 2F5 y el C-terminal que contiene los epítipos 4E10/10E8/Z13, mientras que el péptido CpreTM contiene este último epítipo (residuos 671-683), que es también fundamental en el proceso de fusión de las membranas víricas y celular, más el segmento Leu684-Val693, que corresponde con el extremo N-terminal de TMD. Para aumentar la solubilidad, el péptido CpreTM contiene tres Lys en el extremo N-terminal y dos en el C-terminal. En el centro del epítipo 4E10, el péptido CpreTM tiene Asp674 en lugar de Asn674 dado que la secuencia nativa de gp41 presenta indistintamente Asn o Asp en esta posición.

Por otra parte, se ha diseñado el péptido TMDp, figura 4.17, que contiene la región TMD completa más dos Lys en el extremo N-terminal y tres en el extremo C-terminal, que se han añadido para aumentar la solubilidad.

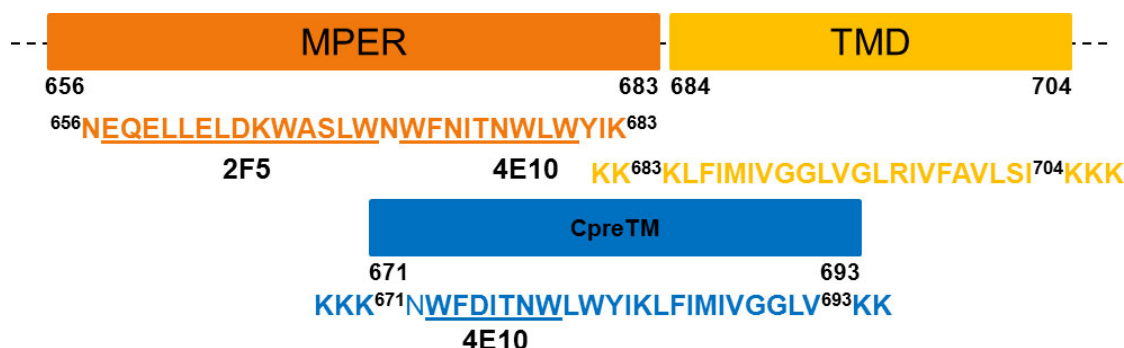
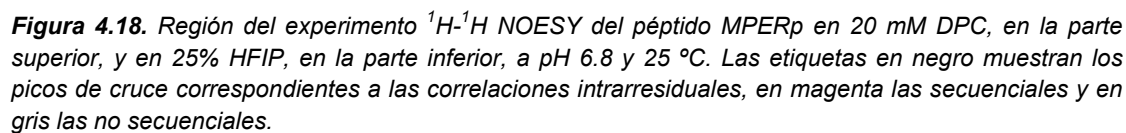
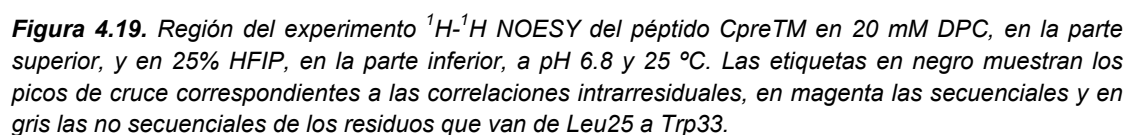


Figura 4.17. Representación esquemática de las regiones MPER y TMD de las que derivan las secuencias de los péptidos MPERp, CpreTM y TMDp de la glicoproteína gp41 del VIH-1. Las zonas subrayadas corresponden con los epítipos de reconocimiento 2F5 y 4E10.





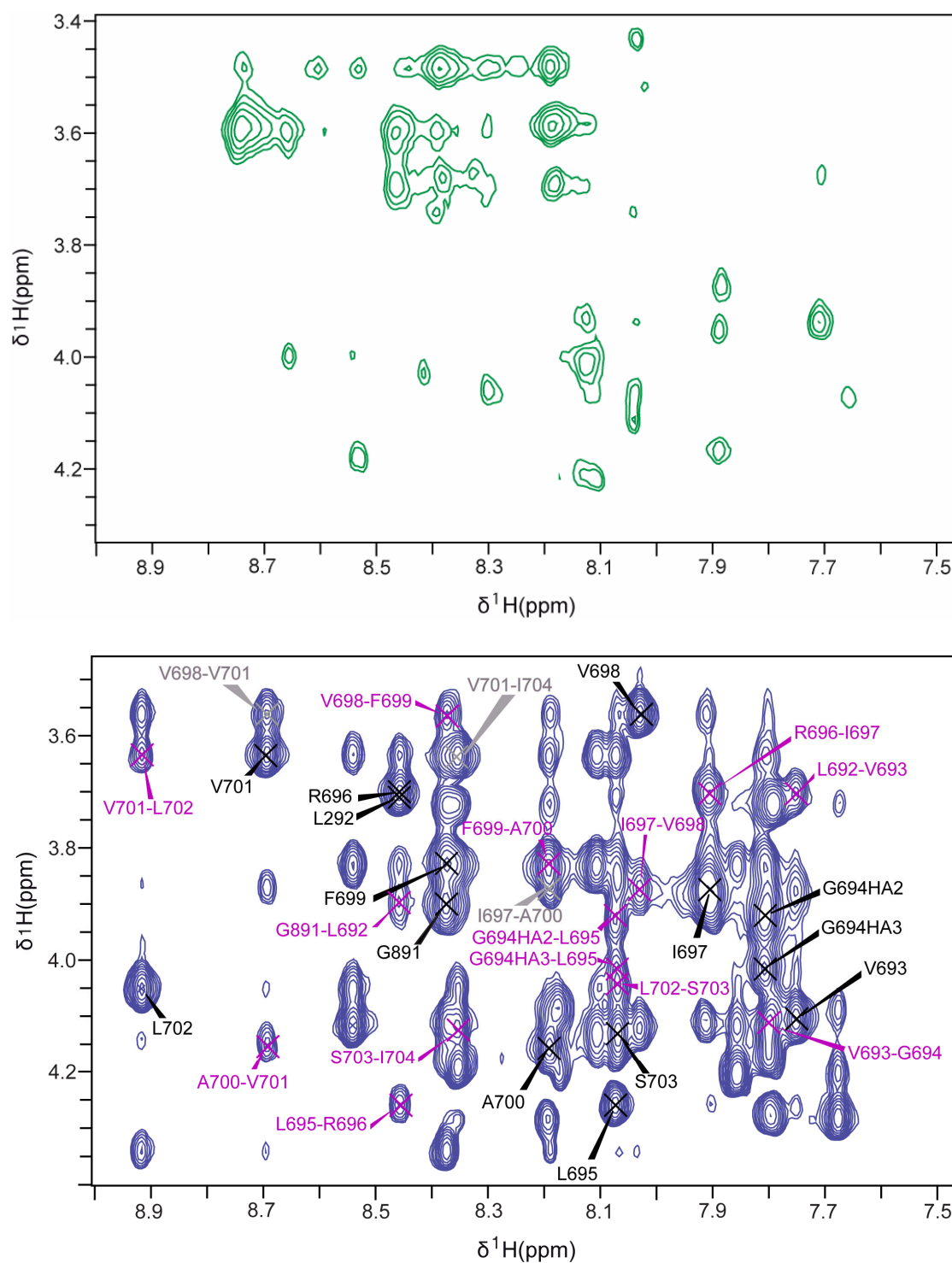


Figura 4.20. Espectros ^1H - ^1H NOESY del péptido TMDp. En la parte superior en 20 mM DPC y en la inferior en 25% HFIP. Las etiquetas en negro muestran los picos de cruce correspondientes a las correlaciones intrarresiduales, en magenta las secuenciales y en gris las no secuenciales.

3.2.2. Espectros de RMN

Los espectros de RMN de los péptidos MPERp, CpreTM y TMDp se han adquirido en 25% HFIP y en 20 mM DPC a 25 °C. La calidad de los espectros, figuras 4.18, 4.19, y 4.20, ha permitido la asignación completa de los desplazamientos químicos de ^1H para los péptidos MPERp y CpreTM tanto en 20 mM DPC como en 25% HFIP y para el péptido TMDp en 25% HFIP (tablas A.5-A.9 del apéndice). La presencia de las micelas de DPC conduce a un ensanchamiento de las señales y una disminución de la relación señal/ruido, que hace más laboriosa la asignación en los péptidos MPERp y CpreTM, y la imposibilita en el péptido TMDp, donde el ensanchamiento es mayor. Para el péptido CpreTM en 25% HFIP fue posible la asignación de los desplazamientos químicos de ^{13}C a partir de los espectros ^1H - ^{13}C HSQC adquiridos en abundancia natural de ^{13}C . Los desplazamientos químicos a pH 7.0 y a una temperatura de 25 °C se han depositado en la base de datos BioMagResBank (<http://www.bmrb.wisc.edu/>) con los códigos de acceso BMRB-19263 (MPERp en 20 mM DPC), BMRB-19262 (MPERp en 25% HFIP), BMRB-19581 (TMDp en 25% HFIP), BMRB-19582 (CpreTM en 25% HFIP) y BMRB-19583 (CpreTM en 20mM DPC).

3.2.3. Análisis cualitativo de la estructura secundaria

La mayoría de los desplazamientos químicos de $^1\text{H}_\alpha$ de los péptidos MPERp, CpreTM y TMDp, tanto en HFIP como en DPC, muestran desviaciones con respecto a los valores característicos de péptidos sin estructura o random coil. En particular, la mayoría de los protones H_α presentan desplazamientos conformacionales, $\Delta\delta_{\text{H}_\alpha}$, negativos y grandes en valor absoluto, figura 4.21. Estos resultados indican claramente que MPERp, CpreTM y TMDp adoptan estructura helicoidal tanto en presencia de HFIP como en micelas de DPC. En el péptido CpreTM algunos residuos presentan valores pequeños en magnitud, tanto positivos como negativos, que podrían indicar una cierta irregularidad en la hélice formada, o deberse a efectos de anisotropía de los anillos aromáticos de los Trp (672, 678 y 680). En el caso del péptido TMDp la magnitud pequeña que presentan los residuos Leu684, Gly690, Gly691 y Leu692 es indicativa de una cierta irregularidad en la hélice formada, ya que el péptido TMDp carece de residuos aromáticos.

Además de los datos de los desplazamientos químicos, en los tres péptidos los NOEs observados apoyan de forma clara la formación de estructura helicoidal. Así, los tres péptidos presentan NOEs secuenciales $d_{\text{NN}(i,i+1)}$ intensos y NOEs no secuenciales y de medio alcance como $d_{\alpha\text{N}(i,i+2)}$, $d_{\alpha\text{N}(i,i+3)}$, $d_{\alpha\text{N}(i,i+4)}$, $d_{\alpha\beta(i,i+3)}$, característicos de hélices α , que pueden observarse en la figura 4.22.

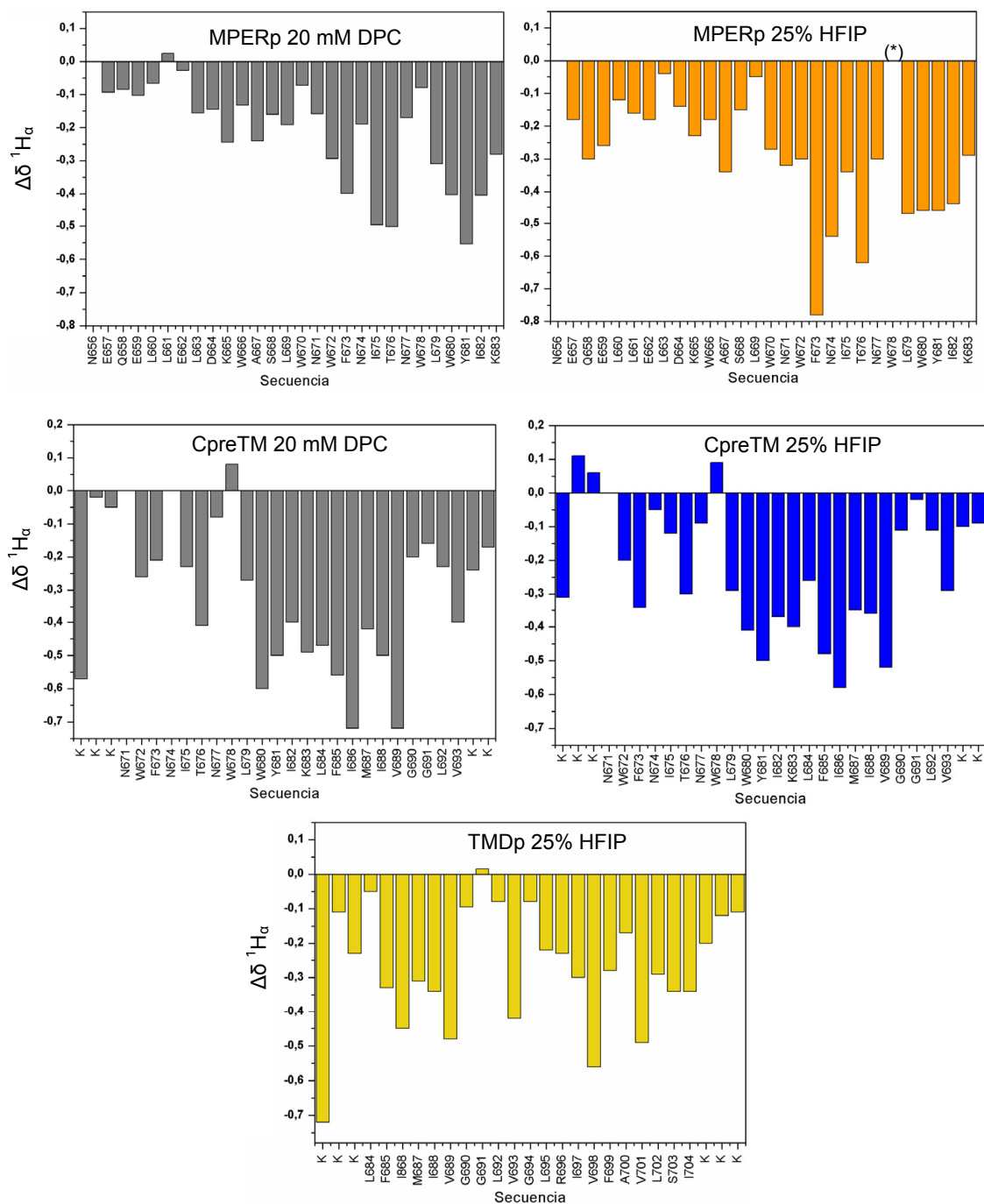


Figura 4.21. Desplazamientos conformacionales de $^1\text{H}_\alpha$ ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm) de los péptidos MPERp, CpreTM y TMDp. A la izquierda, están representados los $\Delta\delta$ en 20 mM DPC y a la derecha, en 25% HFIP excepto en el caso de TMD que sólo se ha determinado en HFIP. Los valores negativos para los $^1\text{H}_\alpha$ indican la formación de hélice α . El símbolo (*) indica que el valor de desplazamiento químico correspondiente no ha podido ser determinado.

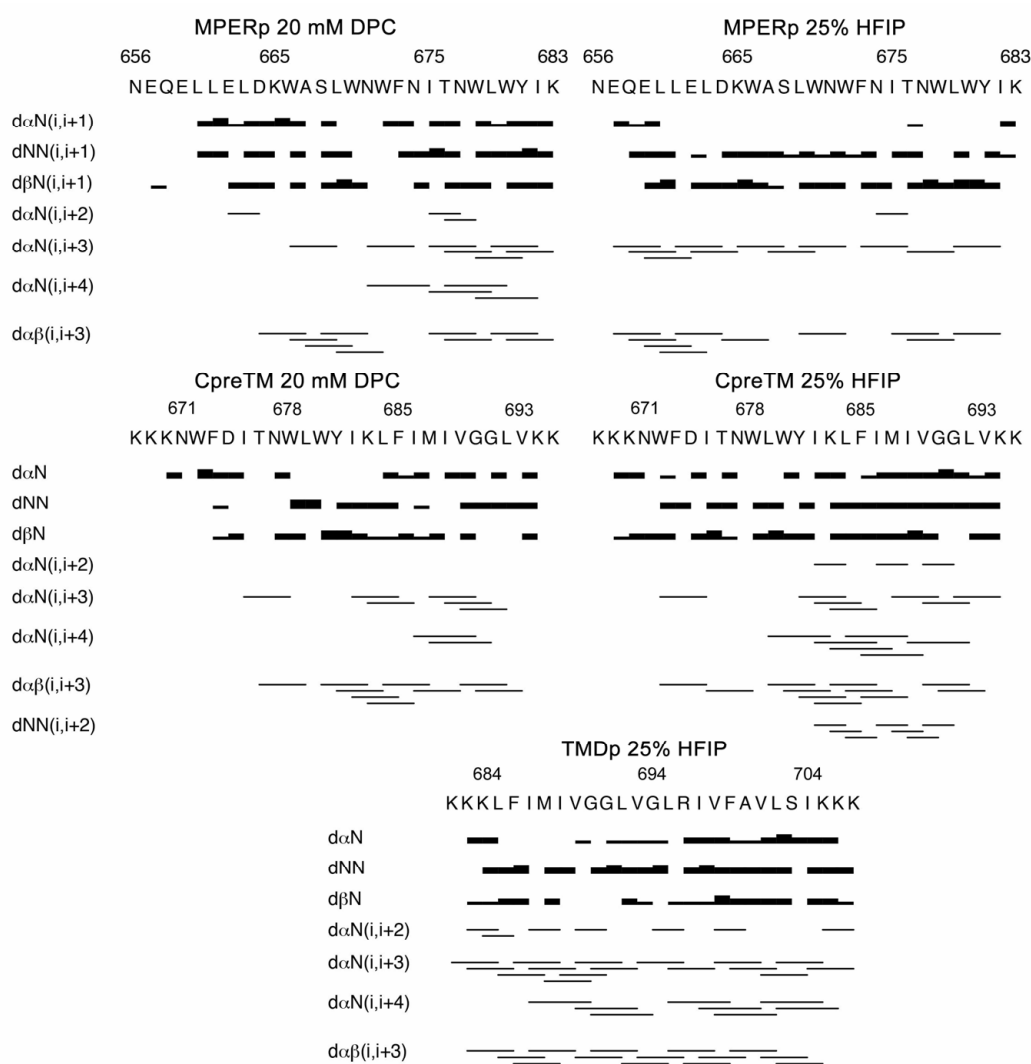


Figura 4.22. NOEs observados en los espectros ^1H - ^1H NOESY de los péptidos MPERp, CpreTM y TMDp en 25 % HFIP y 20 mM DPC, excepto en el caso de TMDp, obtenidos en un espectrómetro AV600 MHz empleando un tiempo de mezcla de 150 ms.

		MPERp		TMDp	CpreTM	
		DPC	HFIP	HFIP	DPC	HFIP
Restricciones de distancia	Intrarresiduales ($i - j = 0$)	154	118	129	126	156
	Secuenciales ($ i - j = 1$)	103	74	99	61	106
	Rango Medio ($1 < i - j < 5$)	115	35	138	112	98
	Número Total	372	227	366	299	360
	Restricciones por residuo	13.3	8.1	13.6	10.7	12.9
Número de restricciones angulares	Ángulos ϕ	26	24	25	23	25
	Ángulos ψ	25	24	25	24	22
	Número Total	51	48	50	47	47
Violaciones máximas por estructura	Distancia (Å)	0.15±0.01	0.02±0.04	0.03	0.05±0.01	0.01±0.01
	Ángulos Diedros (°)	3.5 ± 0.1	0.05±0.08	0.5	1.2±0.5	0.03±0.04
Valores de energía de las estructuras	Función blanco de CYANA	0.35±0.01	5.4E-3±9.7E-3	2.95E-3±2.4E-5	7.4E-2±1.8E-2	6.04E-4±6E-4
	Energía AMBER (kcal/mol)	- 1187	- 1203	-824	-839	-861
	Energía de Van der Waals (kcal/mol)	- 195	- 191	-157	-164	-180
	Energía Electrostática (kcal/mol)	- 2247	- 2232	-1434	-1850	-2006
Desviaciones de la geometría ideal	Longitud de enlaces (Å)	0.014	0.015	0.014	0.015	0.015
	Ángulos de enlace (°)	1.8	1.7	1.7	1.9	1.8
RMSD por pares	Átomos del esqueleto	0.4 ± 0.4	1.6 ± 0.6	1.0±0.3	1.1 ± 0.5	1.8±0.6
	Átomos pesados	1.1 ± 0.3	2.4 ± 0.7	1.9±0.3	1.9 ± 0.7	2.8±0.6
Mapa de Ramachandran (%)	Residuos en regiones más favorables	99.6	98.5	100	95.4	96.7
	Residuos en regiones permitidas adicionales	0.4	1.5	0	4.6	3.3
	Residuos en regiones permitidas generosamente	0.0	0.0	0	0.0	0.0
	Residuos en regiones no permitidas	0.0	0.0	0	0.0	0.0

Tabla 4.3. Datos estadísticos de las 20 estructuras de menor energía calculadas para los péptidos MPERp, TMDp y CpreTM.

3.2.4. Estructuras tridimensionales de los péptidos

Para visualizar las estructuras de los péptidos y las posibles diferencias entre las estructuras adoptadas en DPC y en HFIP, se calcularon las estructuras de los tres péptidos en HFIP, y de MPERp y CpreTM en DPC siguiendo el protocolo descrito en la sección 2.6. Las familias de las 20 estructuras calculadas, para cada péptido y condición, cumplen las correspondientes restricciones experimentales, tanto angulares como de distancia y son energéticamente favorables, ya que presentan un bajo valor de función blanco o valores negativos en la energía obtenida con AMBER, tabla 4.3. Las estructuras se encuentran bien definidas, como se deduce de los bajos valores de RMSD, tabla 4.3, indicativos de la pequeña desviación que se da en las distintas estructuras calculadas para cada péptido, y por tanto, de la buena convergencia de las mismas, como puede apreciarse en las figuras 4.23, 4.27, y 4.29. Además, prácticamente todos los ángulos se encuentran en las zonas más favorables del mapa de Ramachandran, tabla 4.3. Estas estructuras se han depositado en la base de datos del PDB con los códigos de acceso 2M8M (MPERp en 25% HFIP), 2M8O (MPERp en 20 mM DPC), 2MG1 (TMDp en 25% HFIP), 2MG2 (CpreTM en 25% HFIP) y 2MG3 (CpreTM en 20 mM DPC).

Todas las estructuras son de tipo helicoidal, tal como se había deducido del análisis cualitativo de desplazamientos químicos y NOEs (sección 3.2.3). En el caso del péptido MPERp, las estructuras calculadas en HFIP y en DPC son muy similares, como puede apreciarse en la superposición de los dos conjuntos de 20 estructuras calculados, figura 4.23. Las diferencias más notables aparecen en los residuos del extremo N- terminal en HFIP, donde la convergencia de las estructuras varía, hecho que también se refleja en un valor de RMSD más alto con respecto a las estructuras calculadas en DPC, y en la posición de los cambios de dirección del eje helicoidal. En HFIP, la hélice se dobla en la posición ⁶⁶⁵KW⁶⁶⁶ formando un ángulo de 40°, figura 4.26. Mientras que en presencia de DPC, figura 4.26, los residuos del extremo N-terminal ⁶⁶¹LEL⁶⁶³ forman una hélice 3_{10} , y tras el siguiente segmento de hélice α hay un giro en el tripéptido ⁶⁷³FNI⁶⁷⁵, formando un ángulo de unos 30°.

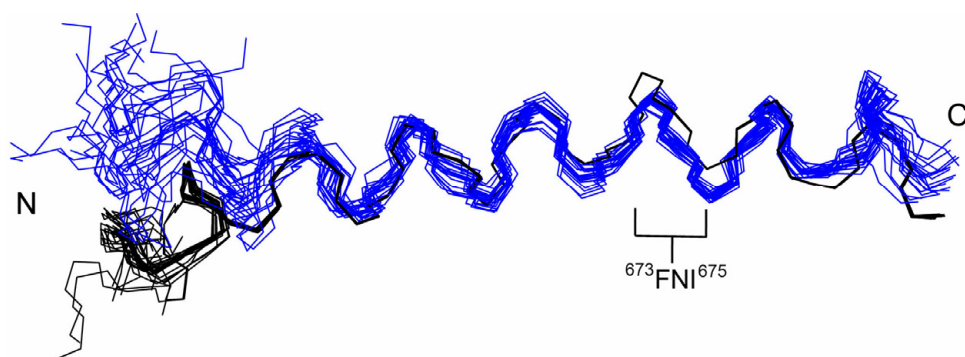


Figura 4.23. Superposición del péptido MPERp, en azul están representadas las 20 estructuras en 25% HFIP y en negro en 20mM DPC. El esqueleto del fragmento Leu663-Tyr681 prácticamente se superpone en las dos familias de estructuras, a excepción del giro $^{673}\text{FNI}^{675}$ que se da en la estructura en DPC.

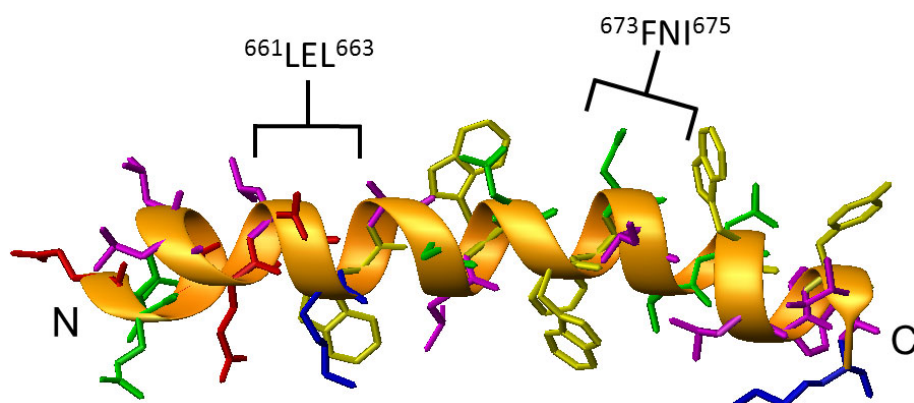


Figura 4.24. Estructura tipo cinta del péptido MPERp en 20 mM DPC donde la hélice α va del residuo Gln658 hasta Lys683. Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu y Ala; en verde a Asn, Gln, Ser, Thr; en rojo a Asp y Glu; en azul a Lys; y en amarillo a Tyr, Phe y Trp.

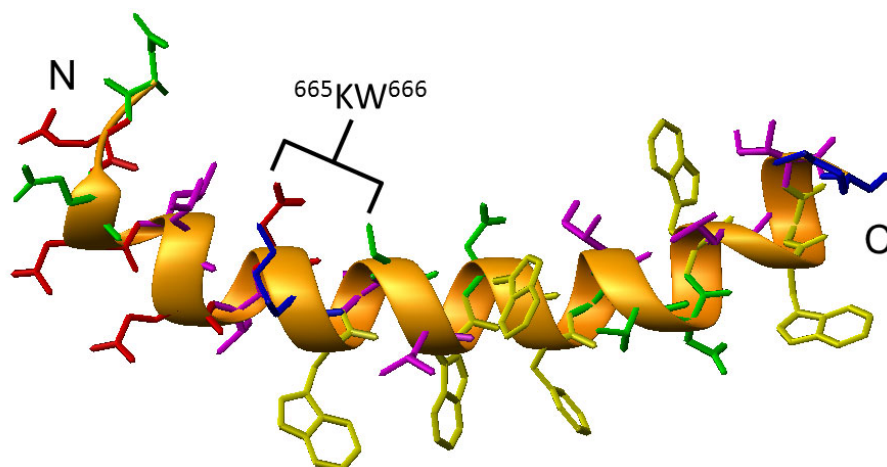


Figura 4.25. Estructura de hélice α del péptido MPERp en 25% HFIP que va del residuo Glu657 a la Lys683. Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu y Ala; en verde a Asn, Gln, Ser, Thr; en rojo a Asp y Glu; en azul a Lys; y en amarillo a Tyr, Phe y Trp.

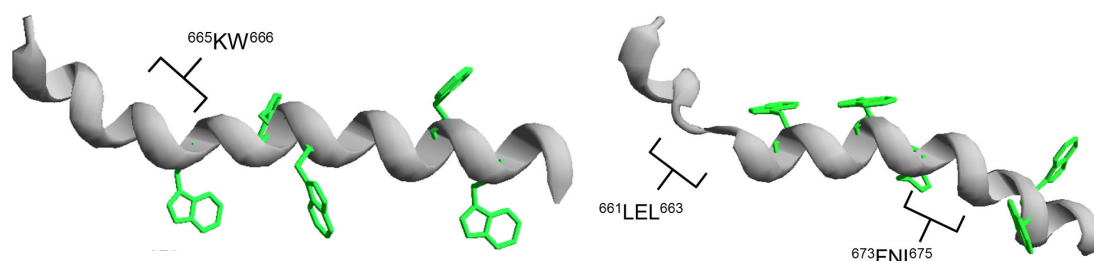


Figura 4.26. A la izquierda estructura en cinta del péptido MPERp en 25% HFIP con el correspondiente giro de la región helicoidal en los residuos ⁶⁶⁵KW⁶⁶⁶. A la derecha, el péptido MPERp en 20 mM DPC donde se observan dos giros en los residuos ⁶⁶¹LEL⁶⁶³ y ⁶⁷³FNI⁶⁷⁵. Las cadenas laterales de los triptófanos están representadas en verde.

En cuanto a la hélice adoptada por TMDp, figuras 4.27 y 4.28, prácticamente abarca todo el péptido, puesto que va de Lys682 hasta Lys706, y presenta un giro en los residuos ⁶⁹¹GLV⁶⁹³.

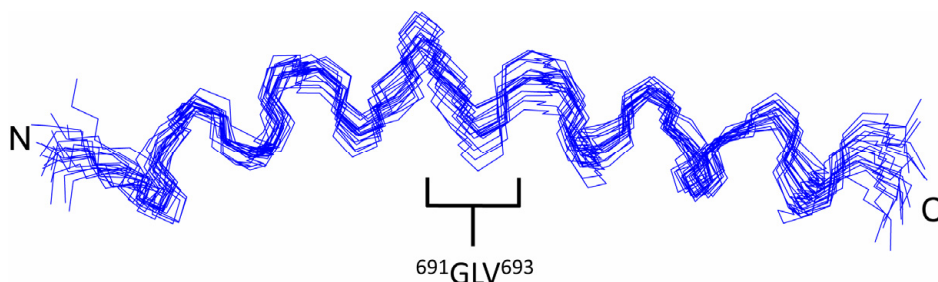


Figura 4.27. Superposición de la familia de 20 estructuras con valor de energía más bajo calculadas para el péptido TMDp en 25% HFIP

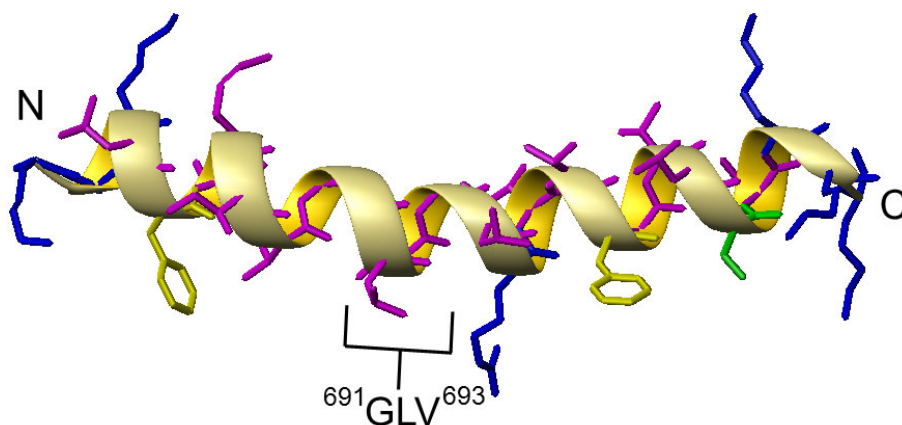


Figura 4.28. Estructura helicoidal tipo cinta del péptido TMDp en 25% HFIP. Las cadenas representadas están magenta representan residuos de Ile, Leu, Met, Val, Gly y Ala; en verde de Ser; en azul de Lys y Arg; en amarillo de Phe.

La superposición de las dos familias de estructuras calculadas para CpreTM en DPC y HFIP de la figura 4.29 muestra que en el fragmento Trp672–Gly691 son prácticamente superponibles y es en los extremos N- y C- terminal donde más varían las estructuras con respecto al eje helicoidal principal. Las diferencias son mayores en el extremo C-terminal de la familia de estructuras en DPC.

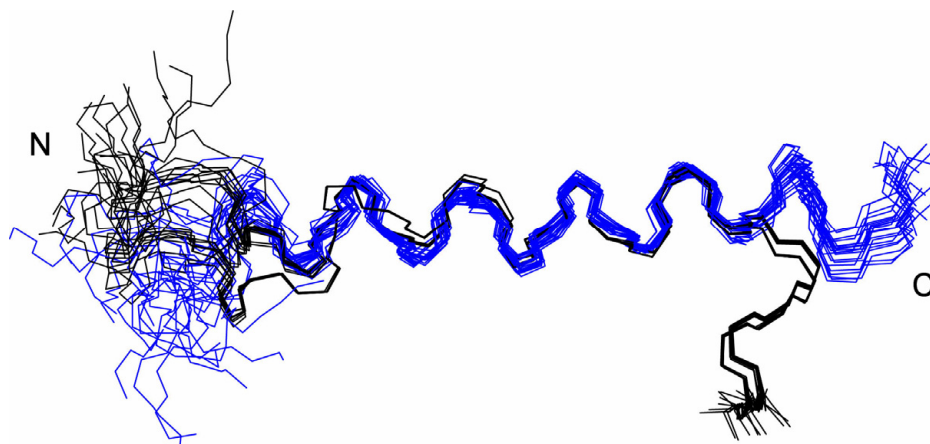


Figura 4.29. Superposición de las familias de estructuras obtenidas para el péptido CpreTM en 25% HFIP, en azul, y 20 mM DPC, en negro.

La estructura de CpreTM en 20 mM DPC, figura 4.30, está formada por tres hélices. Los tres primeros residuos, que corresponden a las tres Lys no nativas del péptido, adoptan una estructura de hélice 3_{10} que va seguida de dos residuos sin estructura, que la separan de una segunda hélice tipo α que va de Ile675 a Val689. En el extremo C-terminal se forma una tercera hélice de tipo 3_{10} de corta longitud, de Gly690 a Lys694.

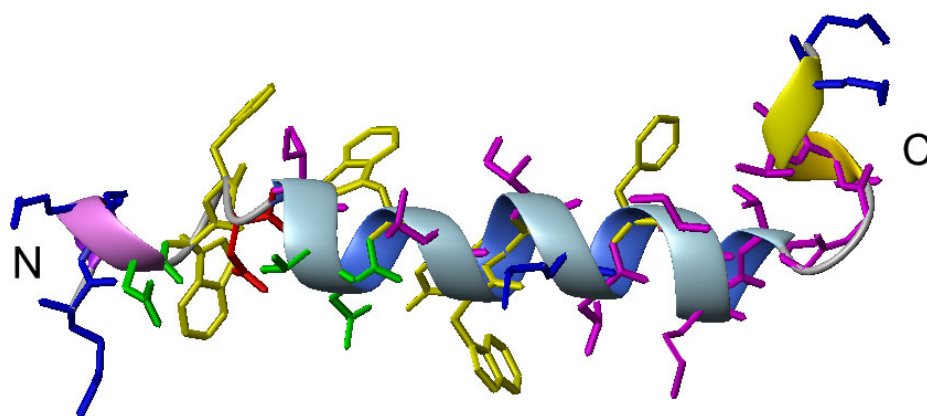


Figura 4.30. Estructura tipo cinta del péptido CpreTM en 20 mM DPC. En magenta primera hélice 3_{10} , en azul región en hélice α y en amarillo segunda hélice 3_{10} . Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu, Met, Val y Gly; en verde a Asn y Thr; en rojo a Asp; en azul a Lys; y en amarillo a Tyr, Phe y Trp.

Los tres primeros residuos, que corresponden a las tres Lys no nativas del péptido adoptan una estructura de hélice que no tiene continuidad con el resto. Al final en el extremo C-terminal, encontramos una tercera hélice que afecta a los últimos cuatro residuos Gly690, Gly691, Leu692, Val693 (sin contar las dos últimas Lys).

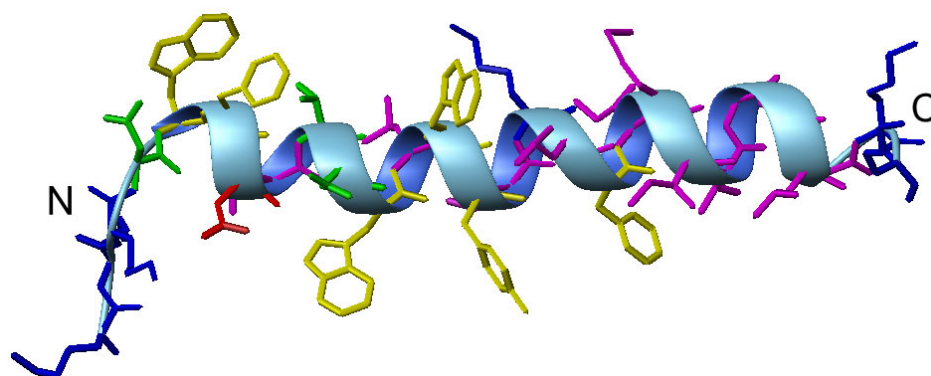


Figura 4.31. Estructura tipo cinta del péptido CpreTM en 25% HFIP. Las cadenas laterales representadas en magenta corresponden a residuos de Ile, Leu, Met, Val y Gly; en verde a Asn y Thr; en rojo a Asp; en azul a Lys; y en amarillo a Tyr, Phe y Trp.

4. DISCUSIÓN

4.1. Estructura de la región FP

Como se expuso en la Introducción (sección 1.5), la región FP de gp41, y en particular, los residuos altamente conservados de su centro, son esenciales en los procesos de fusión del virus (Weissenhorn et al. 1999; Nieva & Agirre 2003). Sin embargo, las estructuras conocidas de alta resolución determinadas para el ectodominio no incluyen la región FP, debido a su reducida solubilidad (Weissenhorn et al. 1997a). Por ello, la estructura que esta región tiene en el péptido híbrido HybK3 resulta de gran interés.

De acuerdo con los datos de RMN, las estructuras adoptadas por la región FP del péptido HybK3 en presencia de 25% HFIP y en 20 mM DPC muestran diferencias estructurales, figura 4.32. En los dos medios, la estructura es helicoidal, pero en HFIP se extiende desde el residuo Ala517 hasta Lys530, mientras que en DPC adopta una hélice α desde Ile515 a Phe522, seguida de una hélice 3_{10} de Leu523 a Ala525, con los residuos ⁵²²FL⁵²³ formando un giro que origina un cambio en la dirección de la hélice α con respecto a la hélice 3_{10} . El giro en ⁵²²FL⁵²³ puede ser debido a las interacciones de los residuos no polares con las micelas de DPC. Como se observa en la figura 4.32, las cadenas laterales en la zona central de la hélice α mantienen prácticamente la misma disposición en los dos medios. Las mayores diferencias se observan en los extremos, en particular en el C-terminal donde los residuos 524-528 no tienen una estructura definida en DPC. Estas diferencias estructurales en función del medio en el que se encuentra el péptido sugieren que la estructura que adoptará la región FP puede verse influenciada por la composición de la membrana, lo que está de acuerdo con los estudios previos de Jaroniec (Jaroniec et al. 2005).

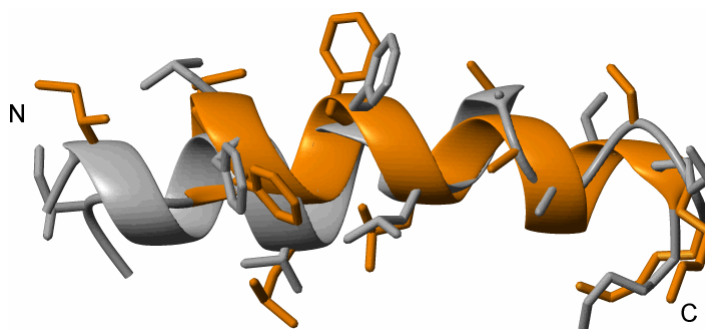


Figura 4.32. Región FP, Gly514 a Lys530 del péptido HybK3. La estructura tipo cinta y cadenas laterales en gris muestran la estructura en 20 mM DPC y en naranja en 25% HFIP.

La estructura resuelta en HFIP para el péptido HybK3 (donde la hélice va de Ala517 a Lys530) tiene una gran similitud con la determinada para el péptido correspondiente a los residuos 512-535 en micelas de SDS (2ARI) (Jaroniec et al. 2005) que muestra una hélice ininterrumpida donde los 12 residuos del segmento Ile515-Ala526 están protegidos del disolvente y embebidos en la micela, mientras que los residuos anterior y posterior, Gly514 y Gly527, se encuentran en la interfase micelar. Por el contrario, la estructura calculada en DPC (donde la región helicoidal va de Ile515 a Ala525) se asemeja más a la estructura del péptido que contiene los residuos 512-535 de gp41 determinada también en micelas de DPC (código pdb 2PJV) (Li & Tamm 2007), el cual muestra una hélice de Ile515 a Ala525, pero con una pequeña variación de la estructura helicoidal en los residuos Leu520-Phe522. Las similitudes encontradas entre las estructuras de la región FP del péptido HybK3 (residuos 514-528) con las estructuras de péptidos de mayor longitud (2ARI y 2PJV) indican que la estructura se mantiene en los péptidos más cortos.

Por otra parte, resulta especialmente llamativo que la región scrFP del péptido scrHybK3 mantenga la capacidad de formación de hélice a pesar del cambio de orden de los residuos conservados centrales del péptido de fusión (de LFLGFLG a FGLLGFL, figura 4.7). En HFIP, la longitud de la hélice de scrFP es aproximadamente igual que la adoptada por la región FP en el péptido HybK3. Evidentemente, la disposición de las cadenas laterales de Phe y Leu es diferente, figura 4.33. Sin embargo, en presencia de 20 mM DPC, la hélice formada por scrFP es significativamente más corta, de Leu521 a Ala526, que la de FP en el péptido HybK3, que va de Ile515 a Ala525, figura 4.34. Este acortamiento de la hélice α puede ser indicativo de la importancia de la conservación del centro de la región FP para su comportamiento estructural. Esto último explicaría que la estructura helicoidal de la región FP de HybK3, cuyo centro está conservado, no se vea tan influenciada por el medio como la región scrFP en scrHybK3. Teniendo en cuenta estos resultados, los datos previos de que incluso mutaciones puntuales relativamente conservativas pueden tener efectos dramáticos en la actividad de fusión (Steinhauer et al. 1995; Delahunty et al. 1996; Pritsker et al. 1999), tienen que explicarse por la importancia concreta de las interacciones en que intervenga el residuo mutado, y no por el efecto sobre la formación de la estructura helicoidal.

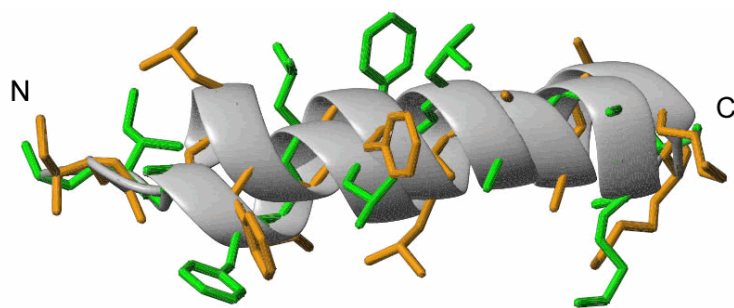


Figura 4.33. Región FP en 25% HFIP. Las cadenas laterales del péptido HybK3 están representadas en naranja y las del péptido scrHybK3 en verde.

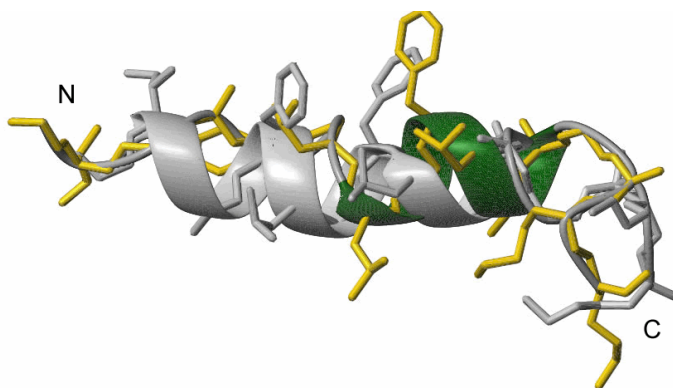


Figura 4.34. Representación de la región FP en 20 mM DPC, donde el péptido HybK3 está representado por la estructura tipo cinta y las cadenas laterales en gris y el péptido scrHybK3 por la cinta en verde y las cadenas laterales en amarillo.

4.2. Estructura de la región MPER e implicaciones para la inmunogenicidad

La importancia de la región MPER viene dada por el hecho de que los epítomos 2F5 y 4E10/Z13 de gp41 están localizados en esta región (Burton et al. 2004; Lorizate et al. 2008; Montero et al. 2008; Sun et al. 2008; Gray et al. 2009). Esto implica que los péptidos que reproduzcan la estructura del dominio inmunogénico MPER pueden, en teoría, ser empleados como vacunas contra el VIH. Aunque ha sido objeto de numerosos estudios, los datos que apoyan la conexión entre la inmunogenicidad y la estructura de péptidos derivados de la región MPER son escasos. Se han estudiado péptidos de distintas longitudes que mostraban indicios de estructuras parciales, del tipo hélice 3_{10} y hélice α (Schibli et al. 2001; Biron et al. 2002; Joyce et al. 2002), pero ninguno de ellos muestra una estructura helicoidal continua como las determinadas aquí para el péptido MPERp.

La hélice α adoptada por el péptido MPERp en presencia de HFIP presenta un pequeño cambio en la dirección de su eje en los residuos $^{665}\text{KW}^{666}$, figuras 4.25 y 4.26, y en micelas de DPC la región N-terminal es de tipo de hélice 3_{10} y tiene un cambio de dirección en los residuos $^{673}\text{FNI}^{675}$, figuras 4.24 y 4.26. Para analizar la relevancia de estas estructuras para la inmunogenicidad del péptido MPERp conviene compararlas con las que los epítomos 2F5 y 4E10 tienen en los complejos con los correspondientes anticuerpos.

Empezando por el epítomo 2F5, figura 4.35, en las estructuras de MPERp en HFIP el tripéptido $^{664}\text{DKW}^{666}$, que comprende el núcleo central del epítomo 2F5, quedaría enterrado por el parátomo de acuerdo a la estructura del complejo. Estos residuos forman un giro β de tipo I con características conformacionales correspondientes a la transición entre estructuras extendidas y helicoidales. En la figura 4.35 se comparan las estructuras adoptadas por los residuos $^{664}\text{DKW}^{666}$ en diferentes péptidos. Las estructuras en DPC de MPERp, así como de péptidos de menor longitud (2PV6), presentan una interacción alquil- π entre las cadenas laterales de Lys665 y Trp666 similar a la existente en el complejo con el anticuerpo (3D0L), figura 4.35. Sin embargo, ninguna de la estructuras de los péptidos aislados reproduce la disposición de la carga negativa de la cadena lateral del Asp664 en el complejo Fab-péptido (3D0L) (Bryson et al. 2009). Por otra parte, hay que destacar que las estructuras de la región MPER presente en los péptidos quimera HybK3 y scrHybK3 (ver también sección 4.4), que prácticamente corresponde al epítomo 2F5 (residuos 655-671), son muy similares a las de esta región en el péptido MPERp, figura 4.16. Este hecho indica que la estructura del epítomo 2F5 es independiente de la presencia del epítomo 4E10.

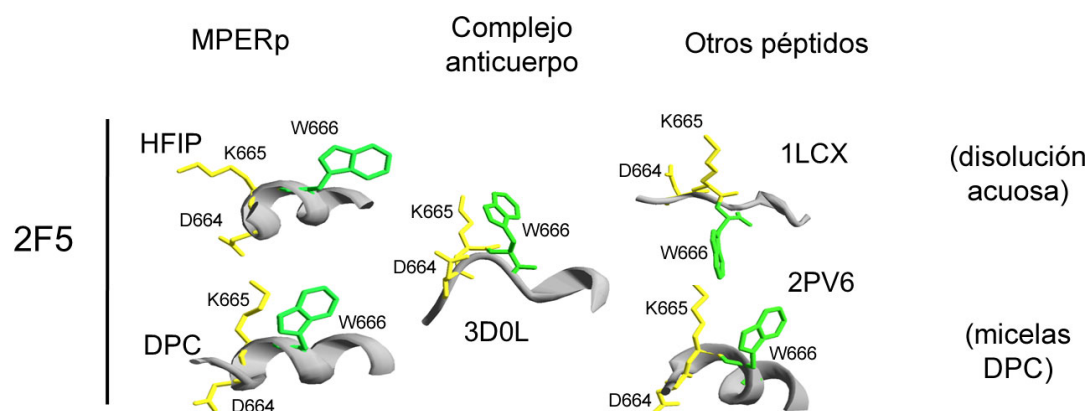


Figura 4.35. Orientación de los residuos que forman el centro ⁶⁶⁴DKW⁶⁶⁶ del epítipo de 2F5. El código de acceso PDB está indicado al lado de cada estructura. DPC y HFIP indican las estructuras del péptido MPERp en 20 mM DPC y 25% HFIP determinadas en este trabajo, respectivamente; 3D0L corresponde a la estructura del complejo Fab-epítipo depositada en el PDB; 1LCX es la estructura en solución del PDB; 2PV6 representa la estructura del PDB en presencia de micelas de DPC.

En el epítipo 4E10, que se encuentra presente en el péptido CpreTM además de en MPERp, las cadenas laterales de los residuos aromáticos Trp672 y Phe673 se orientan de forma similar en los dos péptidos en HFIP y de forma diferente en DPC, figura 4.36. La disposición que estas cadenas laterales tienen en el péptido MPERp en DPC debida a la formación del giro ⁶⁷³FNI⁶⁷⁵, figura 4.26, es muy similar a la que presentan en un péptido en su complejo con el anticuerpo Fab4E10 (Cardoso et al. 2007), figura 4.36. De hecho, el péptido MPERp reproduce mejor la disposición que estas cadenas laterales tienen en los complejos péptido/anticuerpo (2FX7) que el péptido CpreTM o que péptidos más cortos que, al igual que CpreTM contienen el epítipo 4E10 completo, pero no el 2F5, como son 2PV6 (Sun et al. 2008) y 1JAV (Schibli et al. 2001). Esto indica que la presencia del epítipo 2F5 completo puede favorecer que la región del epítipo 4E10 adopte la conformación adecuada para la unión al anticuerpo, a diferencia de lo observado para el epítipo 2F5.

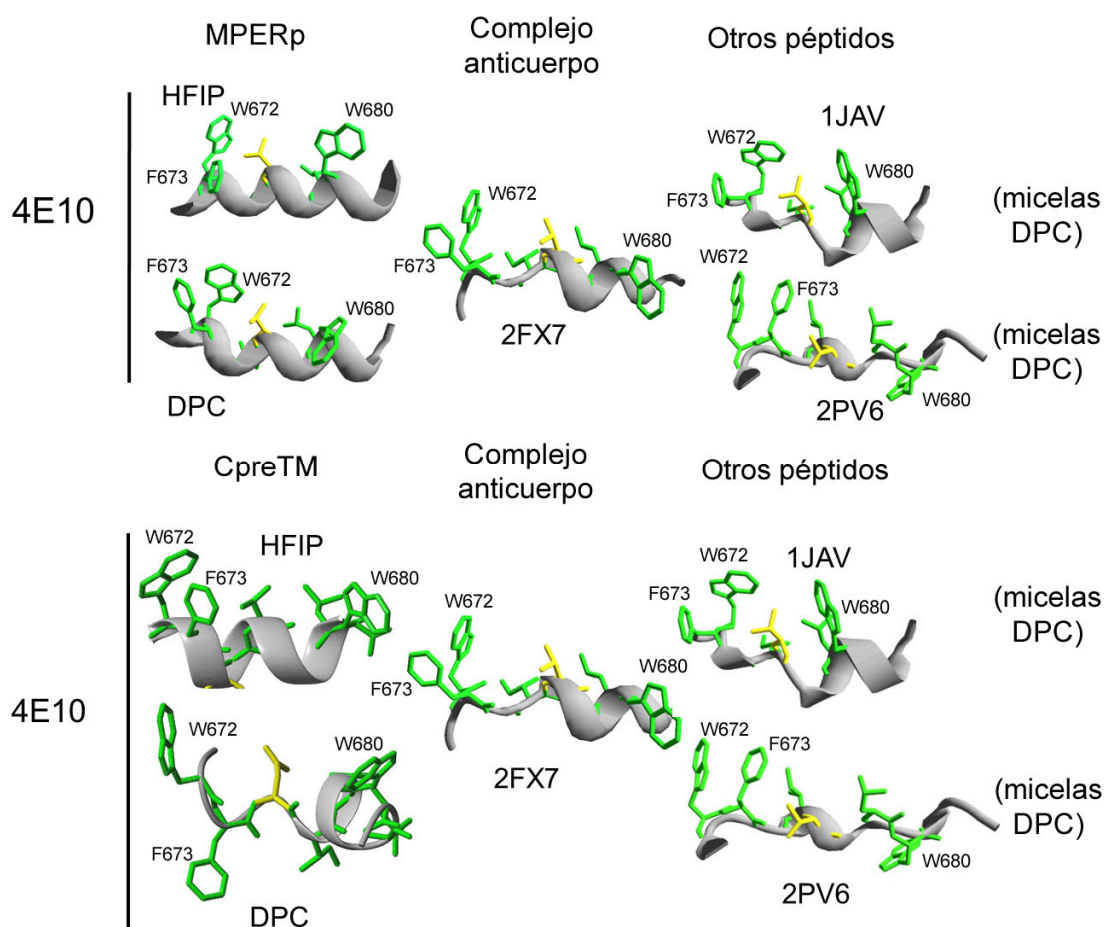


Figura 4.36. Orientación de los residuos que forman el centro $^{673}\text{FNI}^{675}$ del epítipo 4E10 en MPERp, en la parte superior y de los residuos que forman el centro $^{673}\text{FDI}^{675}$ del epítipo 4E10 en CpreTM, en la parte inferior. Se indican los códigos de acceso PDB de las distintas estructuras. DPC y HFIP representan las estructuras del péptido MPERp y CpreTM en 20 mM DPC y 25% HFIP determinadas en este trabajo, respectivamente; 2FX7 corresponde a la estructura del complejo Fab-epítipo; 1JAV y 2PV6 indican estructuras del PDB en presencia de micelas de DPC.

4.3. Interacción entre las regiones FP y MPER

El análisis estructural de los péptidos híbridos HybK3 y scrHybK3 nos permite discutir las posibles interacciones entre la región FP y la parte de la región de MPER (⁶⁵⁶KNEQELLELDKWASLWN⁶⁷¹) que coincide prácticamente con el sitio de reconocimiento del epítipo 2F5. Las dos regiones adoptan estructuras helicoidales. Las características de las hélices formadas por la región FP en los dos péptidos se han expuesto en la sección 3.1. Las hélices adoptadas por la región del epítipo 2F5 son prácticamente iguales en los dos péptidos, e independientes de las condiciones de HFIP o micelas de DPC, como se puede observar en la figura 4.37. La disposición de las cadenas laterales también es bastante parecida. Aunque son bastante similares, las hélices son ligeramente más largas en HFIP (de Glu657 a Asp664) que en DPC (Gln657 a Ser 668). A pesar de la similitud en la estructura global de esta región en los dos péptidos híbridos, el péptido HybK3 presenta un giro en ⁶⁶⁶WA⁶⁶⁷, que se da tanto en DPC como en HFIP y que corresponde con el que se observa en el péptido MPERp en 25% HFIP, figura 4.38, que no aparece en el caso del péptido scrHybK3.

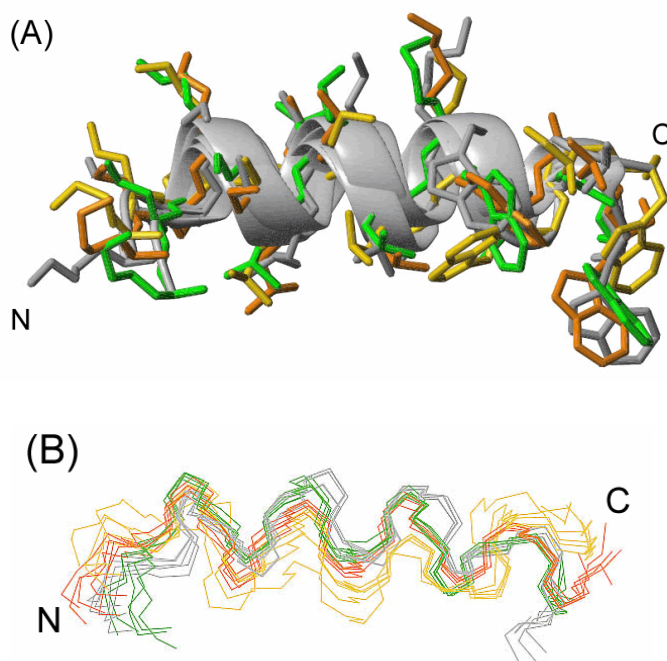


Figura 4.37. Representación de la región MPER de los péptidos HybK3 y scrHybK3. En naranja HybK3 en HFIP y en gris HybK3 en DPC; en verde scrHybK3 en HFIP y en amarillo scrHybK3 en DPC. (A) Superposición de la estructura tipo cinta con las cadenas laterales. (B) Superposición del esqueleto de las cuatro estructuras más favorables energéticamente de cada uno de los péptidos en HFIP y DPC.

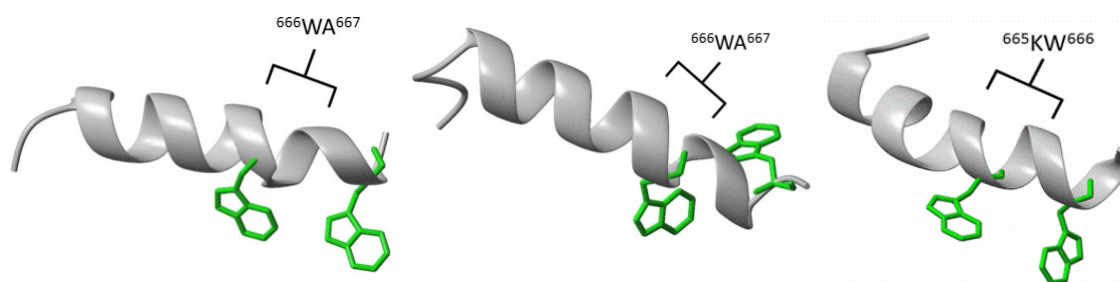


Figura 4.38. Representación del epítipo de 2F5. A la izquierda región MPER de HybK3 en 25% HFIP y en el centro en 20 mM DPC. A la derecha epítipo de 2F5 del péptido MPERp en 25% HFIP.

Las distintas estructuras obtenidas para ambos péptidos muestran variabilidad en la disposición de una hélice con respecto a la otra en ambas condiciones, como reflejan los elevados valores globales de RMSD para las estructuras (en torno a 6-7 Å, mientras que los valores de RMSD de ambas regiones de manera independiente son bastante bajos, aproximadamente 1 Å, tabla 4.2) y la gran variabilidad de los ángulos entre las hélices de FP y MPER. Esta flexibilidad en la posición relativa de las hélices, consecuencia de la falta de NOEs entre los residuos de la región FP y los de MPER, impide confirmar la hipótesis inicial de que las interacciones entre las regiones FP y MPER ayuden a mantener la estructura del tallo de gp41 antes de producirse la fusión de las membranas (Lorizate et al. 2006a; de la Arada et al. 2009), figura 4.4, aunque tampoco la descartan. El modelo simplificado utilizado, el péptido HybK3, sustituye los 127 residuos que conectan las regiones FP y MPER en la proteína nativa gp41 por tan sólo dos Lys y un ácido 6-amino hexanoico, figura 4.7. Por tanto, es posible que la interacción FP/MPER requiera de las interacciones entre residuos no incluidos en el modelo. De todos modos, los resultados están a favor de los resultados de Tamm y colaboradores que indican que los ectodominios y los péptidos de fusión se pliegan de forma independiente (Tamm et al. 2002).

4.4. Estructura de la región TMD

La región TMD de gp41 está insertada en la membrana conectando la región extracelular, el ectodominio que contiene las regiones FP y MPER, con la región intracelular, el dominio citosólico, CT. Con anterioridad al estudio descrito en este capítulo, no se había resuelto la estructura de ninguna construcción de gp41 que contuviera la región TMD debido a los problemas asociados con la expresión y purificación de proteínas hidrófobas. La incorporación de residuos cargados positivamente en los extremos N- y C-terminales con el fin de aumentar la solubilidad del péptido ha sido satisfactoria. Así, el péptido TMDp ha sido suficientemente soluble para poder adquirir espectros de RMN tanto en 25 % HFIP como en 20 mM DPC, aunque sólo los primeros han podido ser completamente asignados. Esto ha permitido la determinación de la estructura tridimensional del péptido TMDp en presencia de HFIP. Esta estructura, que es la primera conocida de la región TMD de gp41, es de tipo helicoidal, lo que coincide con los estudios de modelización computacional anteriores que sugerían que la región TMD adopta hélice α (Gallagher et al. 1989; Domene & Illingworth 2012). La estructura presenta un giro en el centro de la hélice, ⁶⁹¹GLV⁶⁹³, figura 4.39. Este giro podría ser importante para la funcionalidad de gp41, actuando como una bisagra que facilita los cambios conformacionales que gp41 sufre durante el proceso de fusión de las membranas.

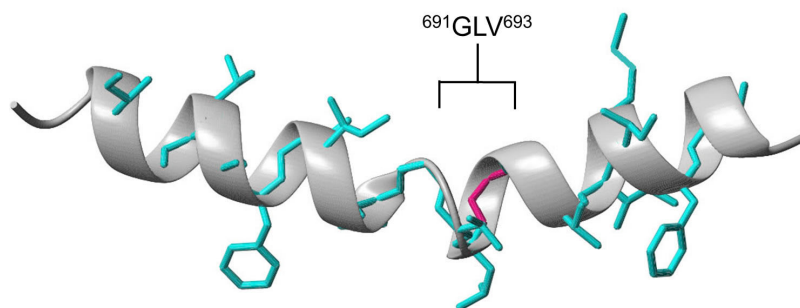


Figura 4.39. Representación tipo cinta del péptido TMD en 25% HFIP. Se indican los residuos del giro que origina un cambio en la dirección de la hélice. Las cadenas laterales están representadas en azul a excepción del residuo que está en el centro del giro, la Leu692, que se muestra en magenta.

4.5. Implicaciones de la estructura de los péptidos en el mecanismo de fusión de las membranas

Los péptidos HybK3, MPERp, CpreTM y TMDp en su conjunto abarcan los dominios MPER y TMD de la proteína gp41, y por pares tienen regiones de secuencia solapantes. Aunque ya se ha comentado la similitud estructural de estas regiones entre algunos de estos pares, como es el caso de HybK3 y MPERp que comparten el segmento correspondiente al epítipo 2F5 (sección 4.2 de esta Discusión), es interesante analizar las similitudes y diferencias estructurales en el conjunto de todos los péptidos. Para este fin, en primer lugar se ha examinado la representación de los $\Delta\delta$ de los protones H_α para los péptidos HybK3, MPERp, CpreTM y TMDp en HFIP, figura 4.40. y en DPC (no mostrada). De forma general, se observa que los perfiles de cada región son iguales independientemente del péptido al que pertenezcan, y sólo se observan variaciones en la magnitud de $\Delta\delta$. Estas variaciones corresponden a residuos localizados en las estructuras helicoidales, lo que se explicaría por la diferente estabilidad de dichas regiones en los distintos péptidos. Además, hay coincidencia en la posición de los valores de menor magnitud (próximos a cero) que en su mayor parte coinciden con los giros observados en las estructuras helicoidales. De forma cualitativa, esto indica que la estructura global adoptada por cada segmento de secuencia es independiente del péptido en el que se encuentre.

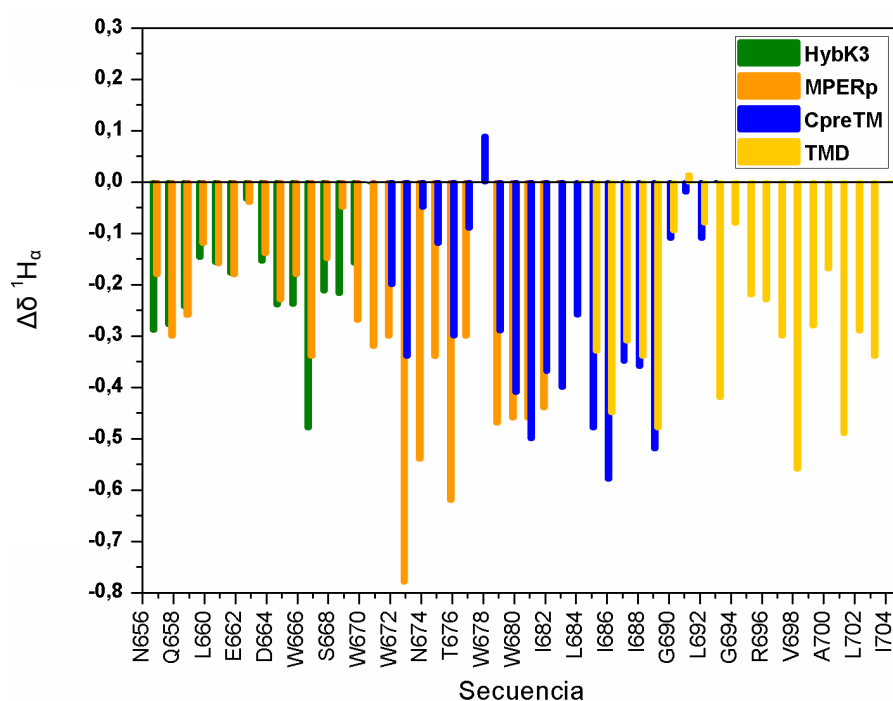


Figura 4.40. Diagrama de barras de los desplazamientos químicos conformacionales de $^1H_\alpha$ ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm) de los péptidos HybK3, MPERp, CpreTM y TMDp en 25% HFIP.

Las analogías estructurales entre los distintos segmentos de las regiones MPER y TMD deducidas a partir de los desplazamientos conformacionales de los H_{α} se evidencian mediante el análisis de la superposición del esqueleto de las estructuras tridimensionales. Así, las estructuras determinadas en este trabajo para los péptidos MPERp y CpreTM en DPC muestran una excelente superposición para el esqueleto de la región Asn671-Thr676, correspondiente al epítipo 4E10, figura 4.41. Sin embargo, presentan diferencias apreciables en la organización de las cadenas laterales, figura 4.42. Estas diferencias podrían explicar la distinta capacidad inmunogénica de los péptidos (datos del grupo de J.L. Nieva), ya que la disposición de las cadenas laterales parece importante para la interacción con los anticuerpos (sección 4.2).

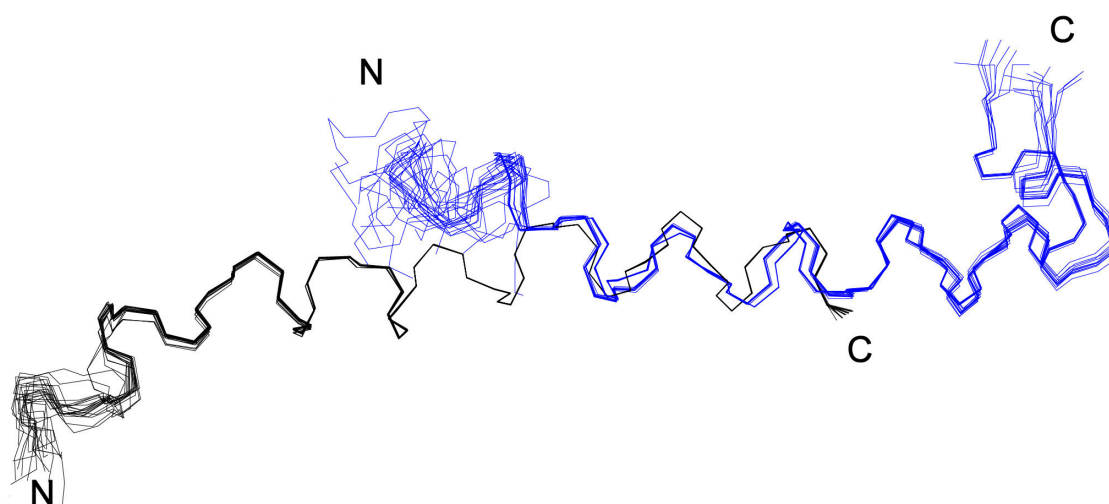


Figura 4.41. Superposición del esqueleto de las 20 estructuras calculadas del péptido MPERp, en negro, y CpreTM, en azul ambos en 20 mM DPC.

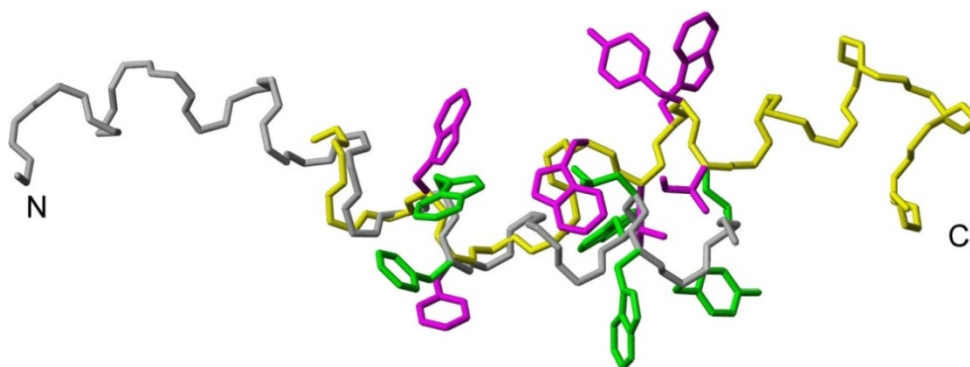


Figura 4.42. Superposición de una de las estructuras del péptido MPERp, con el esqueleto gris y las cadenas laterales en verde, y CpreTM con el esqueleto en amarillo y las cadenas laterales en morado.

Por otra parte, una característica común a las estructuras helicoidales de todos estos péptidos es la presencia de giros que cambian la dirección del eje helicoidal, uno en la región MPER (alrededor de N/D⁶⁷⁴), presente en los péptidos MPERp y CpreTM, y otro en la región TMD (alrededor de G⁶⁹¹), en los péptidos CpreTM y TMDp. Estos giros pueden ser relevantes para entender el mecanismo de fusión de las membranas.

Empezando por el giro presente en la región MPER, presenta diferencias estructurales dependiendo de la longitud del péptido empleado y del medio en el que se encuentra, figura 4.43. La diferencia más notable se encuentra en el ángulo entre las regiones que preceden y siguen al giro ⁶⁷³FNI⁶⁷⁵ en el péptido MPERp. A medida que aumenta la estructura helicoidal en la región C-terminal disminuye el ángulo entre los segmentos anteriores y posteriores a este giro. Estos resultados están a favor de la hipótesis de que el giro ⁶⁷³FNI⁶⁷⁵ desempeña un papel clave en la inducción de la fusión de las membranas.

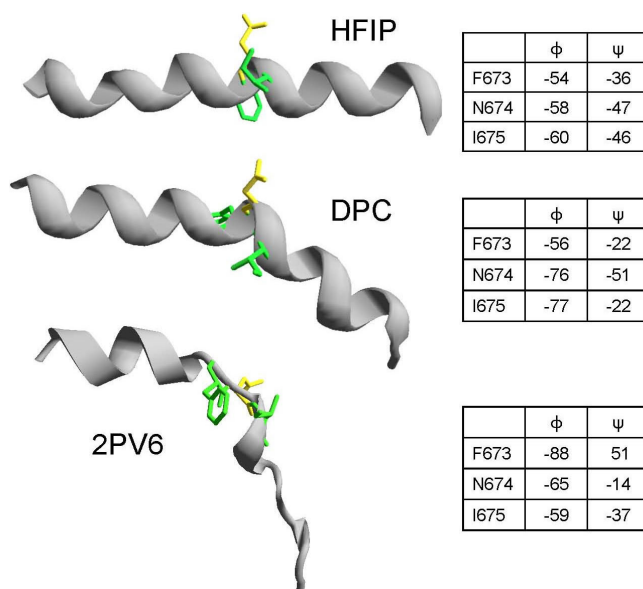


Figura 4.43. A la izquierda representación de los péptidos MPERp en HFIP y DPC, y 2PV6 con el giro que sufren en ⁶⁷³FNI⁶⁷⁵, en verde están representadas las cadenas laterales del Phe673 e Ile675 y en amarillo de Asn674. A la derecha, valores de los ángulos Φ y ψ para cada uno de estos residuos.

En el péptido CpreTM, este giro que muestra MPERp en DPC, ⁶⁷³FNI⁶⁷⁵, figura 4.43, está presente tanto en DPC, ⁶⁷³FDI⁶⁷⁵, como en presencia de HFIP, aunque desplazado, ⁶⁷⁴DIT⁶⁷⁶, figura 4.44. A pesar de esta pequeña variación en la posición del giro, las cadenas laterales de los residuos aromáticos que flanquean al mismo están colocadas hacia el mismo lado, por lo que podrían interactuar con la membrana. Se ha argumentado que el giro ⁶⁷³FNI⁶⁷⁵, presente en la estructura 2PV6, figura 4.43, podría permitir el movimiento independiente de los extremos N- terminal y C-terminal, de manera que este último quedaría inmerso en la interfase de la membrana (Sun et al. 2008).

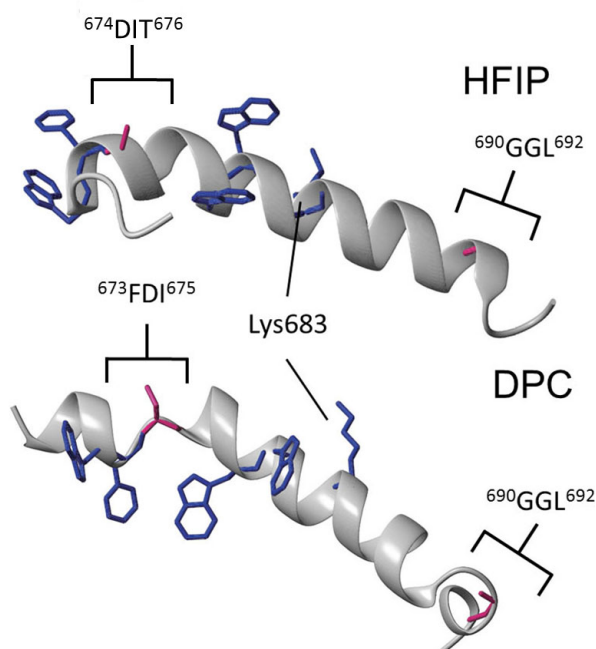


Figura 4.44. Péptido CpreTM en HFIP y DPC. Se indican los giros presentes en las estructuras, así como la Lys683 que marca el final de la región MPER, y el comienzo de la región TMD, inmersa en la membrana plasmática. En magenta están representados los residuos Asp674 y Gly691 que están en el centro de los giros.

A partir de todos estos datos estructurales, podríamos proponer un posible modelo de fusión de las membranas donde el giro $^{673}\text{FNI}^{675}$ presente en la estructura de MPERp en DPC, figura 4.43, o el equivalente $^{673}\text{FDI}^{675}$ presente en CpreTM, figura 4.44, podría ser el elemento gracias al cual la glicoproteína gp41 pudiese conectar el barril de seis hélices 6-HB con la membrana viral curvada que se da durante el intermedio de fusión, figuras 4.45 y 4.46. Este giro haría posible la inserción más profunda de la Lys683 en la membrana, lo que facilitaría la interacción de los residuos aromáticos del fragmento 675-683 con la interfase celular para la formación del poro. A continuación, se produciría el enderezamiento del codo dándose una estructura similar a la resuelta en 25% HFIP. Suponemos que este último movimiento tendría como resultado la extensión del paquete 6-HB en la membrana, que a su vez podría ser acoplado a la ampliación del poro de fusión. Los estudios de mutagénesis de gp41 apoyan el papel del centro del área 671-677 en la expansión del poro (Muñoz-Barroso et al. 1999).

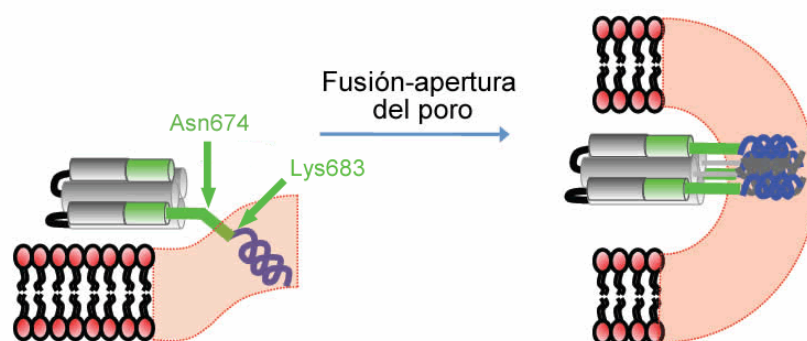


Figura 4.45. Modelo del papel funcional de las estructuras resueltas en este estudio en DPC y HFIP. A la izquierda la hélice torcida podría ser la que une el paquete de seis hélices (6-HB) con alguna curvatura presente en la membrana viral durante el intermedio de fusión. A la derecha transición de MPER a una hélice recta complementaria y trasducción de energía conformacional en la apertura del poro de fusión (Melikyan 2008).

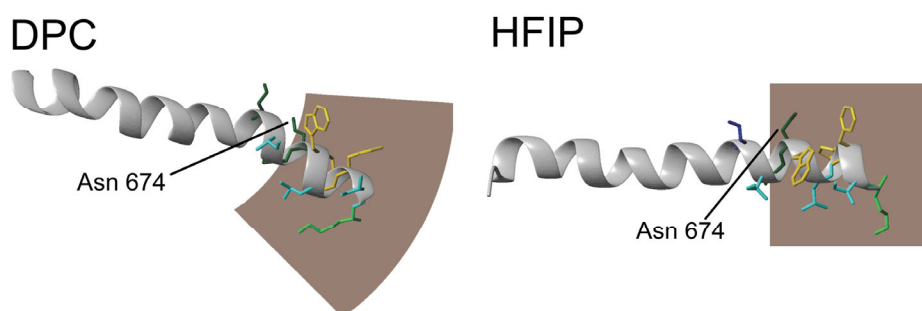


Figura 4.46. Modelo de las estructuras resueltas del péptido MPERp en DPC y HFIP y el supuesto cambio estructural que sufren según van interaccionando con la membrana que está representada en marrón. El residuo señalado Asn674 indica el residuo central del giro $^{673}\text{FNI}^{675}$ en la estructura en DPC.

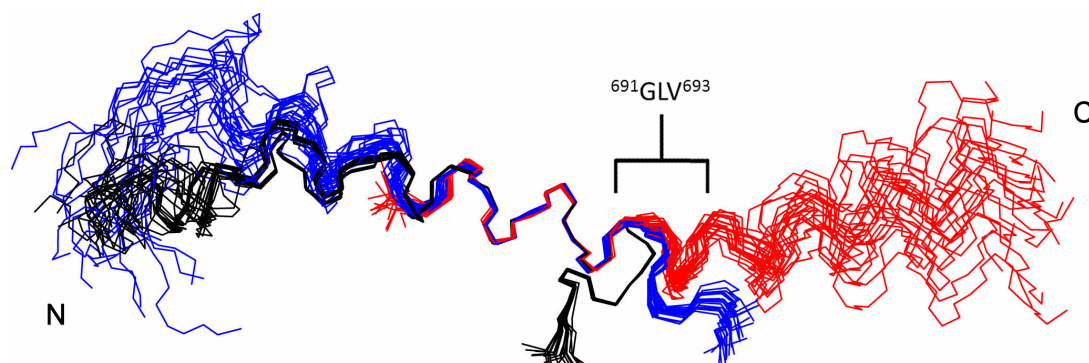


Figura 4.47. A la izquierda CpreTM, en azul en 25% HFIP y negro en 20 mM DPC y a la derecha, en rojo, TMDp en 25% HFIP. Se observa con claridad el intervalo Asn671-Lys683 que se superpone en ambos péptidos y tras el cual se produce la desviación de la hélice a partir del residuo de Gly691.

Finalmente, la superposición de las estructuras de CpreTM y TMDp de la figura 4.47, nos permite observar con mayor claridad el segundo giro presente en estos péptidos, que se produce en la región ⁶⁹⁰GGLV⁶⁹³. Como en el caso anterior, los datos estructurales determinados en este trabajo, nos permiten plantear un modelo de acción para MPER-TMD donde este giro debe tener un papel determinante. Así, este modelo sugiere que MPER-TMD pueda tener dos funciones en la fusión de las membranas. Por una parte que se comporte como palanca para curvar las membranas tal como se visualiza en la figura 4.45. Por otra, que actúe como sacacorchos haciendo que las cadenas alifáticas queden expuestas a la membrana de la célula que se acerca, figura 4.46, en el momento de la prolongación de 6-HB, que como se ha mencionado previamente se postula que actúa como ejecutor de fuerza. La secuencia Gly-Gly es el pivote sobre el que giraría el sacacorchos y es donde la región de la TMD empezaría a disponerse más o menos perpendicular al plano de la membrana.

5. CONCLUSIONES

Empleando la RMN como herramienta se han determinado las estructuras de péptidos correspondientes a diversas regiones de la proteína gp41 de VIH, que juega un papel esencial en la infectividad del virus. Las estructuras se han estudiado en dos medios que mimetizan los entornos de membrana, en presencia de HFIP y en micelas de DPC. La información obtenida contribuye a la comprensión de las interacciones péptido/anticuerpo, y del proceso de fusión de membranas.

En cuanto al medio utilizado para la determinación estructural, las estructuras globales adoptadas en presencia de HFIP y en micelas de DPC son similares, de tipo helicoidal en todos los casos, pero se observan pequeños cambios a nivel local. De forma general, la longitud de las hélices es mayor en HFIP que en DPC. De acuerdo con los resultados en los péptidos quimera HybK3 y scrHybK3, la región FP es más sensible al medio en el que se encuentra y por tanto a la composición de la membrana.

La conservación del centro de la región FP, aunque no es determinante para la estructura tridimensional que adoptan los péptidos scrHybK3 y HybK3, podría afectar a la conformación de la región MPER, como sugiere la ausencia del giro ⁶⁶⁶WA⁶⁶⁷ en la estructura de scrHybK3.

Las similitudes encontradas en las estructuras determinadas para la región MPER (región 2F5 en los péptidos HybK3 y scrHybK3, completa en el péptido MPERp, y región 4E10 en el péptido CpreTM), indican que la selección de los límites de la secuencia no es determinante para mantener la formación de hélice tanto en presencia de 20 mM DPC como en 25% HFIP.

Todas las estructuras determinadas presentan puntos donde se produce una desviación del eje helicoidal principal, cuyo ángulo de desviación varía, y cuya posición se desplaza ligeramente (1-2 residuos) dependiendo de las condiciones del medio. Estas secuencias, muchas de las cuales contienen residuos Gly, podrían actuar como bisagras que permiten a la proteína gp41 adoptar las distintas conformaciones necesarias para cada etapa durante los procesos de fusión de las membranas.

En concreto, la estructura helicoidal que muestra la región MPER, tanto en HybK3 como en MPERp y CpreTM, sufre pequeños cambios conformacionales que le deberían permitir adoptar estructuras energéticamente más favorables en su interacción con la membrana y que se cree que servirían de motor para impulsar la fusión entre las membranas.

De acuerdo a los datos obtenidos en ensayos inmunológicos por nuestro colaborador J.L. Nieva se puede sugerir que las estructuras adoptadas por estos péptidos pueden ser las responsables de su capacidad inmunogénica. Por ello, estos péptidos pueden constituir un punto de partida para el desarrollo de vacunas eficaces frente al VIH.

6. BIBLIOGRAFÍA

- Ardelt W, Mikulski SM, Shogen K (1991) Amino acid sequence of an anti-tumor protein from *Rana pipiens* oocytes and early embryos. Homology to pancreatic ribonucleases. The Journal of biological chemistry 266, 245-251.
- Aue WP, Bartholdi D, Ernst RR (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. . Journal of Chemical Physics 64, 2229-2246.
- Barbato G, Bianchi E, Ingallinella P et al. (2003) Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. Journal of molecular biology 330, 1101-1115.
- Biron Z, Khare S, Samson AO et al. (2002) A monomeric 3(10)-helix is formed in water by a 13-residue peptide representing the neutralizing determinant of HIV-1 on gp41. Biochemistry 41, 12687-12696.
- Bosch ML, Earl PL, Fargnoli K et al. (1989) Identification of the fusion peptide of primate immunodeficiency viruses. Science 244, 694-697.
- Brasseur R, Vandenbranden M, Cornet B et al. (1990) Orientation into the lipid bilayer of an asymmetric amphipathic helical peptide located at the N-terminus of viral fusion proteins. Biochimica et biophysica acta 1029, 267-273.
- Braunschweiler L, Ernst RR (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. Journal of Magnetic Resonance 53, 521-528.
- Brunel FM, Zwick MB, Cardoso RM et al. (2006) Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. Journal of virology 80, 1680-1687.
- Bryson S, Julien JP, Hynes RC et al. (2009) Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. Journal of virology 83, 11862-11875.
- Burton DR (2002) Antibodies, viruses and vaccines. Nature reviews Immunology 2, 706-713.
- Burton DR, Desrosiers RC, Doms RW et al. (2004) HIV vaccine design and the neutralizing antibody problem. Nature immunology 5, 233-236.
- Buzon V, Natrajan G, Schibli D et al. (2010) Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. PLoS pathogens 6, e1000880.
- Buzon V, Padros E, Cladera J (2005) Interaction of fusion peptides from HIV gp41 with membranes: a time-resolved membrane binding, lipid mixing, and structural study. Biochemistry 44, 13354-13364.
- Cardoso RM, Brunel FM, Ferguson S et al. (2007) Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10. Journal of molecular biology 365, 1533-1544.
- Case DA, Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S. and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.
- Cavanagh J, Fairbrother WJ, Palmer III AG et al. (1996) Protein NMR spectroscopy: Principles & Practice. Academic Press.
- Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. Journal of biomolecular NMR 13, 289-302.
- Chan DC, Fass D, Berger JM et al. (1997) Core structure of gp41 from the HIV envelope glycoprotein. Cell 89, 263-273.
- Chan DC, Kim PS (1998) HIV entry and its inhibition. Cell 93, 681-684.

- de la Arada I, Julien JP, de la Torre BG et al. (2009) Structural constraints imposed by the conserved fusion peptide on the HIV-1 gp41 epitope recognized by the broadly neutralizing antibody 2F5. *The journal of physical chemistry B* 113, 13626-13637.
- Delahunty MD, Rhee I, Freed EO et al. (1996) Mutational analysis of the fusion peptide of the human immunodeficiency virus type 1: identification of critical glycine residues. *Virology* 218, 94-102.
- Domene C, Illingworth CJ (2012) Effects of point mutations in pVHL on the binding of HIF-1 α . *Proteins* 80, 733-746.
- Duerst R, Verbach A (1965) Accurate NMR temperature Measurements. *Review of Scientific Instruments* 36, 1896-1897.
- Durell SR, Martin I, Ruyschaert JM et al. (1997) What studies of fusion peptides tell us about viral envelope glycoprotein-mediated membrane fusion (review). *Molecular membrane biology* 14, 97-112.
- Eckert DM, Kim PS (2001) Mechanisms of viral membrane fusion and its inhibition. *Annual review of biochemistry* 70, 777-810.
- Epand RM (2003) Fusion peptides and the mechanism of viral fusion. *Biochimica et biophysica acta* 1614, 116-121.
- Freed EO, Delwart EL, Buchsacher GL, Jr. et al. (1992) A mutation in the human immunodeficiency virus type 1 transmembrane glycoprotein gp41 dominantly interferes with fusion and infectivity. *Proceedings of the National Academy of Sciences of the United States of America* 89, 70-74.
- Frey G, Chen J, Rits-Volloch S et al. (2010) Distinct conformational states of HIV-1 gp41 are recognized by neutralizing and non-neutralizing antibodies. *Nature structural & molecular biology* 17, 1486-1491.
- Gabuzda DH, Lever A, Terwilliger E et al. (1992) Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins. *Journal of virology* 66, 3306-3315.
- Gallaher WR (1987) Detection of a fusion peptide sequence in the transmembrane protein of human immunodeficiency virus. *Cell* 50, 327-328.
- Gallaher WR, Ball JM, Garry RF et al. (1989) A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS research and human retroviruses* 5, 431-440.
- Gallo SA, Finnegan CM, Viard M et al. (2003) The HIV Env-mediated fusion reaction. *Biochimica et biophysica acta* 1614, 36-50.
- Goddard TD, Kneller DG (2005) *Sparky 3*, University of California, San Francisco.
- Gray ES, Madiga MC, Moore PL et al. (2009) Broad neutralization of human immunodeficiency virus type 1 mediated by plasma antibodies against the gp41 membrane proximal external region. *Journal of virology* 83, 11265-11274.
- Güntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353-378.
- Huang J, Ofek G, Laub L et al. (2012) Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* 491, 406-412.
- Jaroniec CP, Kaufman JD, Stahl SJ et al. (2005) Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain. *Biochemistry* 44, 16167-16180.
- Jeener J, Meier BH, Bachmann P et al. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *The Journal of chemical physics* 71, 4546-4553.
- Jiang S, Lin K, Strick N et al. (1993a) HIV-1 inhibition by a peptide. *Nature* 365, 113.
- Jiang S, Lin K, Strick N et al. (1993b) Inhibition of HIV-1 infection by a fusion domain binding peptide from the HIV-1 envelope glycoprotein GP41. *Biochemical and biophysical research communications* 195, 533-538.
- Joyce JG, Hurni WM, Bogusky MJ et al. (2002) Enhancement of α -helicity in the HIV-1 inhibitory peptide DP178 leads to an increased affinity for human monoclonal antibody 2F5 but does not elicit neutralizing responses in vitro. Implications for vaccine design. *The Journal of biological chemistry* 277, 45811-45820.

- Julien JP, Bryson S, Nieva JL et al. (2008) Structural details of HIV-1 recognition by the broadly neutralizing monoclonal antibody 2F5: epitope conformation, antigen-recognition loop mobility, and anion-binding site. *Journal of molecular biology* 384, 377-392.
- Karlsson Hedestam GB, Fouchier RA, Phogat S et al. (2008) The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nature reviews Microbiology* 6, 143-155.
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* 14, 51-55, 29-32.
- Kumar A, Ernst RR, Wuthrich K (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and biophysical research communications* 95, 1-6.
- Laskowski RA, Rullmannn JA, MacArthur MW et al. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR* 8, 477-486.
- Li Y, Tamm LK (2007) Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. *Biophysical journal* 93, 876-885.
- Lin X, Derdeyn CA, Blumenthal R et al. (2003) Progressive truncations C terminal to the membrane-spanning domain of simian immunodeficiency virus Env reduce fusogenicity and increase concentration dependence of Env for fusion. *Journal of virology* 77, 7067-7077.
- Lorizate M, de la Arada I, Huarte N et al. (2006a) Structural analysis and assembly of the HIV-1 Gp41 amino-terminal fusion peptide and the pretransmembrane amphipathic-at-interface sequence. *Biochemistry* 45, 14337-14346.
- Lorizate M, Gomara MJ, de la Torre BG et al. (2006b) Membrane-transferring sequences of the HIV-1 Gp41 ectodomain assemble into an immunogenic complex. *Journal of molecular biology* 360, 45-55.
- Lorizate M, Huarte N, Saez-Cirion A et al. (2008) Interfacial pre-transmembrane domains in viral proteins promoting membrane fusion and fission. *Biochimica et biophysica acta* 1778, 1624-1639.
- Macura S, Ernst RR (1980) Elucidation of cross relaxation in liquids by two-dimensional N.M.R. spectroscopy. *Molecular Physics* 41, 95-117.
- Martin I, Defrise-Quertain F, Decroly E et al. (1993) Orientation and structure of the NH2-terminal HIV-1 gp41 peptide in fused and aggregated liposomes. *Biochimica et biophysica acta* 1145, 124-133.
- Martin I, Ruysschaert JM (1995) Lysophosphatidylcholine inhibits vesicles fusion induced by the NH2-terminal extremity of SIV/HIV fusogenic proteins. *Biochimica et biophysica acta* 1240, 95-100.
- Martin I, Schaal H, Scheid A et al. (1996) Lipid membrane fusion induced by the human immunodeficiency virus type 1 gp41 N-terminal extremity is determined by its orientation in the lipid bilayer. *Journal of virology* 70, 298-304.
- Melikyan GB (2008) Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology* 5, 111.
- Miyauchi K, Komano J, Yokomaku Y et al. (2005) Role of the specific amino acid sequence of the membrane-spanning domain of human immunodeficiency virus type 1 in membrane fusion. *Journal of virology* 79, 4720-4729.
- Montero M, van Houten NE, Wang X et al. (2008) The membrane-proximal external region of the human immunodeficiency virus type 1 envelope: dominant site of antibody neutralization and target for vaccine design. *Microbiology and molecular biology reviews* : MMBR 72, 54-84, table of contents.
- Muñoz-Barroso I, Salzwedel K, Hunter E et al. (1999) Role of the membrane-proximal domain in the initial stages of human immunodeficiency virus type 1 envelope glycoprotein-mediated membrane fusion. *Journal of virology* 73, 6089-6092.

- Muster T, Steindl F, Purtscher M et al. (1993) A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *Journal of virology* 67, 6642-6647.
- Nelson JD, Brunel FM, Jensen R et al. (2007) An affinity-enhanced neutralizing antibody against the membrane-proximal external region of human immunodeficiency virus type 1 gp41 recognizes an epitope between those of 2F5 and 4E10. *Journal of virology* 81, 4033-4043.
- Nieva JL, Agirre A (2003) Are fusion peptides a good model to study viral cell fusion? *Biochimica et biophysica acta* 1614, 104-115.
- Nieva JL, Nir S, Muga A et al. (1994) Interaction of the HIV-1 fusion peptide with phospholipid vesicles: different structural requirements for fusion and leakage. *Biochemistry* 33, 3201-3209.
- Ofek G, Tang M, Sambor A et al. (2004) Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *Journal of virology* 78, 10724-10737.
- Pai EF, Klein MH, Chong P et al. (2000) Fab'-epitope complex from the HIV-1 cross-neutralizing monoclonal antibody 2F5. U.S. patent WO-00/61618.
- Pejchal R, Gach JS, Brunel FM et al. (2009) A conformational switch in human immunodeficiency virus gp41 revealed by the structures of overlapping epitopes recognized by neutralizing antibodies. *Journal of virology* 83, 8451-8462.
- Pereira FB, Goni FM, Muga A et al. (1997a) Permeabilization and fusion of uncharged lipid vesicles induced by the HIV-1 fusion peptide adopting an extended conformation: dose and sequence effects. *Biophysical journal* 73, 1977-1986.
- Pereira FB, Goni FM, Nieva JL (1995) Liposome destabilization induced by the HIV-1 fusion peptide effect of a single amino acid substitution. *FEBS letters* 362, 243-246.
- Pereira FB, Goni FM, Nieva JL (1997b) Membrane fusion induced by the HIV type 1 fusion peptide: modulation by factors affecting glycoprotein 41 activity and potential anti-HIV compounds. *AIDS research and human retroviruses* 13, 1203-1211.
- Poumbourios P, el Ahmar W, McPhee DA et al. (1995) Determinants of human immunodeficiency virus type 1 envelope glycoprotein oligomeric structure. *Journal of virology* 69, 1209-1218.
- Pritsker M, Rucker J, Hoffman TL et al. (1999) Effect of nonpolar substitutions of the conserved Phe11 in the fusion peptide of HIV-1 gp41 on its function, structure, and organization in membranes. *Biochemistry* 38, 11359-11371.
- Purtscher M, Trkola A, Gruber G et al. (1994) A broadly neutralizing human monoclonal antibody against gp41 of human immunodeficiency virus type 1. *AIDS research and human retroviruses* 10, 1651-1658.
- Roux KH, Taylor KA (2007) AIDS virus envelope spike structure. *Current opinion in structural biology* 17, 244-252.
- Salzwedel K, Johnston PB, Roberts SJ et al. (1993) Expression and characterization of glycopospholipid-anchored human immunodeficiency virus type 1 envelope glycoproteins. *Journal of virology* 67, 5279-5288.
- Salzwedel K, West JT, Hunter E (1999) A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *Journal of virology* 73, 2469-2480.
- Sattentau QJ, Zolla-Pazner S, Poignard P (1995) Epitope exposure on functional, oligomeric HIV-1 gp41 molecules. *Virology* 206, 713-717.
- Schaal H, Klein M, Gehrman P et al. (1995) Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. *Journal of virology* 69, 3308-3314.
- Schibli DJ, Montelaro RC, Vogel HJ (2001) The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles. *Biochemistry* 40, 9570-9578.

- Steinhauer DA, Wharton SA, Skehel JJ et al. (1995) Studies of the membrane fusion activities of fusion peptide mutants of influenza virus hemagglutinin. *Journal of virology* 69, 6643-6651.
- Suarez T, Gallaher WR, Agirre A et al. (2000a) Membrane interface-interacting sequences within the ectodomain of the human immunodeficiency virus type 1 envelope glycoprotein: putative role during viral fusion. *Journal of virology* 74, 8038-8047.
- Suarez T, Nir S, Goni FM et al. (2000b) The pre-transmembrane region of the human immunodeficiency virus type-1 glycoprotein: a novel fusogenic sequence. *FEBS letters* 477, 145-149.
- Sun ZY, Oh KJ, Kim M et al. (2008) HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane. *Immunity* 28, 52-63.
- Tamm LK, Han X, Li Y et al. (2002) Structure and function of membrane fusion peptides. *Biopolymers* 66, 249-260.
- Tan K, Liu J, Wang J et al. (1997) Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12303-12308.
- UNAIDS (2009) "Report on the global aids epidemic". United Nations programme on Acquired Immuno Deficiency Syndrome.
- Vranken WF, Boucher W, Stevens TJ et al. (2005) The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* 59, 687-696.
- Weiss CD, White JM (1993) Characterization of stable Chinese hamster ovary cells expressing wild-type, secreted, and glycosylphosphatidylinositol-anchored human immunodeficiency virus type 1 envelope glycoprotein. *Journal of virology* 67, 7060-7066.
- Weissenhorn W, Calder LJ, Dessen A et al. (1997a) Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 gp41 ectodomain expressed in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 94, 6065-6069.
- Weissenhorn W, Dessen A, Calder LJ et al. (1999) Structural basis for membrane fusion by enveloped viruses. *Molecular membrane biology* 16, 3-9.
- Weissenhorn W, Dessen A, Harrison SC et al. (1997b) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387, 426-430.
- Welman M, Lemay G, Cohen EA (2007) Role of envelope processing and gp41 membrane spanning domain in the formation of human immunodeficiency virus type 1 (HIV-1) fusion-competent envelope glycoprotein complex. *Virus research* 124, 103-112.
- Wild CT, Shugars DC, Greenwell TK et al. (1994) Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proceedings of the National Academy of Sciences of the United States of America* 91, 9770-9774.
- Zanetti G, Briggs JA, Grunewald K et al. (2006) Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS pathogens* 2, e83.
- Zheng Z, Yang R, Bodner ML et al. (2006) Conformational flexibility and strand arrangements of the membrane-associated HIV fusion peptide trimer probed by solid-state NMR spectroscopy. *Biochemistry* 45, 12960-12975.
- Zhu P, Liu J, Bess J, Jr. et al. (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441, 847-852.
- Zhu Z, Qin HR, Chen W et al. (2011) Cross-reactive HIV-1-neutralizing human monoclonal antibodies identified from a patient with 2F5-like antibodies. *Journal of virology* 85, 11401-11408.
- Zwick MB, Labrijn AF, Wang M et al. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *Journal of virology* 75, 10892-10905.

Capítulo V

Estructura tridimensional del zimógeno

FLG de Onconasa® activable por la proteasa
del VIH-1

Capítulo V:

Estructura tridimensional del zimógeno FLG de Onconasa® activable por la proteasa del VIH-1

1. INTRODUCCIÓN

1.1. Ribonucleasas y Onconasa®

Como es bien sabido, el RNA es una molécula esencial para la vida que conocemos ya que traduce la información almacenada en el DNA y la convierte en proteínas que son las responsables últimas de la función biológica. Hay dos clases de enzimas que controlan la síntesis y degradación del RNA: las RNA polimerasas, y las RNA despolimerasas, respectivamente. Estas últimas se conocen comúnmente como ribonucleasas (RNasas).

La ribonucleasa mejor conocida es la RNasa A de páncreas bovino (Barnard 1969). Se trata de una proteína ampliamente caracterizada y utilizada como sistema modelo desde los años 30 del siglo pasado. Ha sido utilizada como sistema de referencia en estudios de plegamiento, estabilidad y química de proteínas, enzimología y evolución molecular. Desde un punto de vista histórico, la RNasa A ha sido sujeto importante de los trabajos de Anfinsen, Moore y Stein quienes obtuvieron el Premio Nobel en Química en 1972, así como los de Merrifield que le valieron el Premio en 1984. La ribonucleasa A continua siendo elegida, hoy en día, como sistema modelo de estudio en muchos laboratorios.

Dentro de la familia de las ribonucleasas encontramos la Onconasa® (ONC), también conocida como proteína P-30 o ranpirnasa (Darzynkiewicz et al. 1988), que se extrae de los ovocitos o embriones tempranos de la *Rana pipiens*.

La Onconasa®, con 104 residuos, 20 residuos menos que la RNasa A, es el miembro más pequeño de la superfamilia de la RNasa A. Su estructura primaria es significativamente diferente a la de la RNasa A ya que sólo comparten el 30% de la secuencia. Con una temperatura de fusión de 90 °C su estabilidad térmica es de unos 30 grados más alta que la mayoría de los miembros de las RNasas (Leland et al. 1998). ONC y RNasa A comparten una topología general característica (Mosimann et al. 1994), figura 5.1. Concretamente, las posiciones de las tres hélices α , los sistemas de láminas β , los residuos del centro activo, las regiones hidrófobas y tres de los cuatro enlaces disulfuro son superponibles. La mayoría de las diferencias estructurales más evidentes, entre Onconasa® y RNasa A, se encuentran en la longitud y composición de los bucles y en las regiones N-terminal y C-terminal.

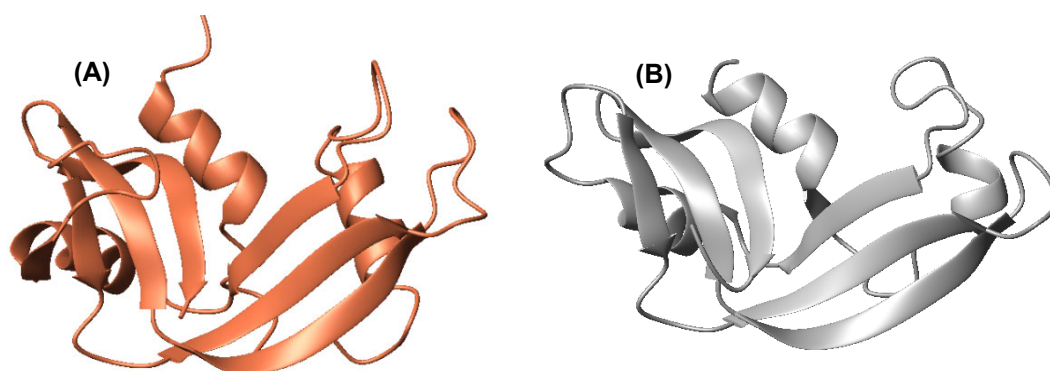


Figura 5.1. Representación gráfica de las estructuras cristalográficas de las proteínas (A) RNasa A, código de acceso pdb 5RSA, y (B) Onconasa®, código de acceso pdb 1ONC. Figuras generadas con el programa MOLMOL (Koradi et al. 1996).

La actividad ribonucleolítica de la ONC es bastante menor que la que presenta su homólogo mamífero, pero a diferencia de éste es mucho más citotóxica (Darzynkiewicz et al. 1988; Ardelt et al. 1991). La alta citotoxicidad de la Onconasa® ha sido atribuida a su rápido proceso de internalización (Saxena et al. 2002; Rodriguez et al. 2007), a su habilidad para evadir al inhibidor de la ribonucleasa presente en el citosol (Boix et al. 1996; Rutkoski & Raines 2008), a su inusual estabilidad conformacional (Notomista et al. 2000; Arnold & Ulbrich-Hofmann 2006) y a su actividad ribonucleolítica (Wu et al. 1993; Lee & Raines 2003; Liao et al. 2003). La citotoxicidad de la Onconasa® le ha permitido llegar a ser la primera ribonucleasa empleada en ensayos clínicos para el tratamiento de distintos tipos de cáncer, habiéndose alcanzado la fase III de los mismos (Mikulski et al. 2002; Costanzi et al. 2005). Además de su citotoxicidad, la Onconasa® posee una demostrada capacidad de inhibición

del virus de inmunodeficiencia humano (VIH) a través de la degradación del RNA vírico (Saxena et al. 1996).

1.2. Zimógenos de RNasas

Un zimógeno o proenzima es un precursor enzimático inactivo. Para activarse, necesita de un cambio bioquímico en su estructura que le lleve a adquirir la conformación activa para poder realizar la catálisis. En la naturaleza, los zimógenos y la proteólisis limitada constituyen los mecanismos básicos para regular la actividad de muchos procesos biológicos. En el caso de los zimógenos, el control reside en la activación de la enzima, en lugar de en su inhibición. El grupo de R. Raines, científico muy conocido en el campo de las RNasas, desarrolló hace un tiempo un zimógeno de RNasa A. Para su diseño unieron los extremos N- y C-terminales de la secuencia nativa a través de un nuevo segmento polipeptídico (Plainkum et al. 2003; Johnson et al. 2006). Además, se generaron nuevos extremos N- y C- terminales diferentes a los nativos mediante un corte en uno de los bucles externos, figura 5.2. El nuevo segmento insertado, abarca y bloquea el sitio activo de la RNasa A, de manera que inhibe su actividad enzimática mientras este “puente” permanezca intacto. En el diseño se incluyó en esta secuencia de unión, un sitio de reconocimiento para la proteasa plasmepsina II. Cuando la proteína así construida se somete a la acción de la proteasa específica, se hidroliza un enlace específico de la secuencia polipeptídica insertada, el centro activo queda accesible y se regenera la actividad catalítica de la RNasa A.

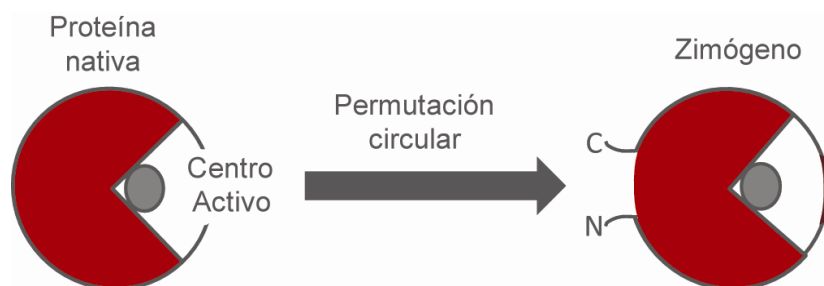


Figura 5.2. Diseño de un zimógeno inactivo a través de una permutación circular de una enzima. Los extremos N- y C- terminales se unen a través de una secuencia polipeptídica insertada (línea gruesa que bloquea el centro activo) y se generan nuevos extremos terminales, mediante un corte en la secuencia.

Sin embargo, a pesar de las expectativas iniciales para el uso de estos zimógenos como tóxicos dirigidos, los zimógenos de la RNasa A no tienen la capacidad de evadir la acción del inhibidor citosólico específico, IR (Dickson et al. 2005). IR es el regulador natural de ribonucleasas en las células de los mamíferos, y se une a la RNasa A y a sus homólogos con una afinidad extremadamente alta ($K_i \sim 10^{-15} \text{M}$). Por ello, terapias basadas en ribonucleasas

deben dirigirse a otras proteínas de la familia que sean capaces por sí mismas de evadir la acción del inhibidor. De lo contrario, esta aproximación está abocada al fracaso. Entre las características que debe cumplir un zimógeno de RNasa destacamos: a) estabilidad elevada; b) buena capacidad de internalización; c) baja actividad ribonucleolítica y citotoxicidad del zimógeno no activado; d) buena activación al reaccionar con el agente activante; e) elevada actividad ribonucleolítica y citotóxica una vez que ha sido activado, figura 5.3.

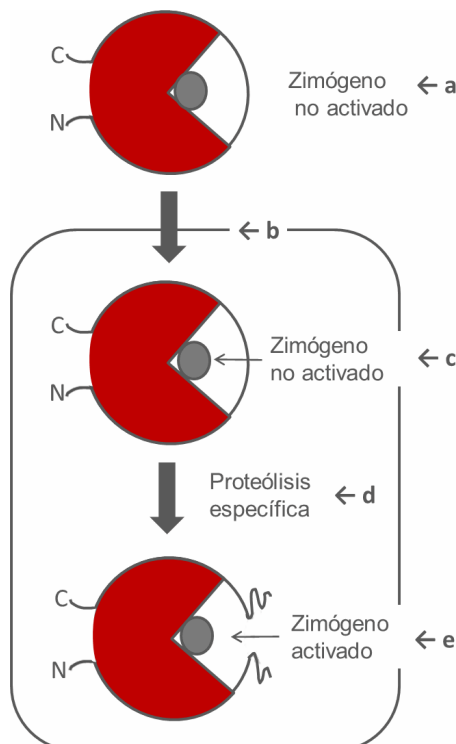


Figura 5.3. Proceso de activación del zimógeno.

Los fines terapéuticos que se pretenden con esta aproximación se basan en que, una vez el zimógeno se ha internalizado, sólo pueda activarse en aquellas células que posean el agente activante para el que han sido diseñados.

En este caso concreto, el zimógeno se activará y podrá desempeñar la actividad ribonucleolítica y citotóxica que produzca la muerte de la célula.

1.3. El zimógeno de la Onconasa®

Como alternativa prometedora al trabajo de Raines comentado en el apartado anterior (Turcotte & Raines 2008), el grupo del Prof. Marc Ribó de la Universitat de Girona propuso que se podrían utilizar zimógenos de la ONC para desarrollar su actividad tóxica de manera dirigida. Esta estrategia tiene en cuenta que la proteína nativa no se une al inhibidor de ribonucleasas, IR, lo que suponía el principal obstáculo para el éxito de los zimógenos de RNasa A. Además, se propuso que podría ser interesante dirigir estas moléculas de manera específica a células infectadas por VIH-1, figura 5.4, de modo que se pudiera desarrollar una terapia alternativa a la utilizada en la actualidad en enfermos de SIDA. Estas células infectadas por el VIH producen la proteasa específica del virus que podría ser usada para la activación del zimógeno y producir así la muerte celular.

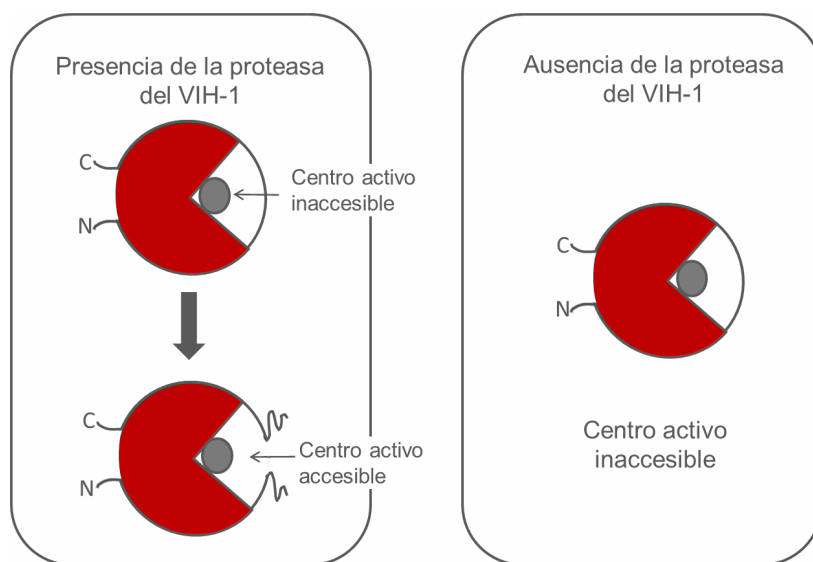


Figura 5.4. Diferencias postuladas para el comportamiento del zimógeno de la Onconasa® si entra en una célula que contiene o no la proteasa del VIH-1. En la figura de la izquierda, la presencia de la proteasa del VIH-1 en una célula infectada, produce el corte del enlace peptídico específico contenido en la secuencia de reconocimiento y el centro activo queda accesible, pudiendo desarrollarse la actividad de la proteína. En la figura de la derecha, correspondiente a una célula no infectada, la proteína permanece inactiva debido a la ausencia de proteasa VIH-1 dentro de la célula.

Es sabido que los primeros medicamentos desarrollados en la lucha contra el SIDA eran análogos de nucleósidos diseñados para obstruir la transcripción inversa, e inhibidores basados en péptidos dirigidos a bloquear la acción de la proteasa del VIH (Tomasselli & Heinrikson 2000; Dauer 2005). La combinación de ambos tipos de medicamentos ha conducido a la exitosa terapia antirretroviral (HAART), la cual tiene como resultado una drástica reducción de la carga del virus con la respectiva reducción de los síntomas asociados con la enfermedad. Sin embargo, la resistencia desarrollada a este tipo de

medicamentos sigue siendo un problema ya que se ha observado que las proteasas mutantes del virus VIH-1, no se inhiben con los fármacos. Este hecho deja abierto un campo de la investigación para el estudio de medicamentos más eficaces en la lucha contra el SIDA (Barre-Sinoussi et al. 1983), y para el desarrollo de nuevas estrategias, en la cuales las posibilidades del virus para eludir la acción de estos medicamentos sean mucho menores. Desde un punto de vista terapéutico la idea del zimógeno que se presenta aquí es importante e interesante pues podría solucionar los conocidos mecanismos de resistencia a los medicamentos que acabamos de comentar. Esta estrategia y la utilización de péptidos inmunogénicos, abordada en el capítulo anterior, se complementan dado que actúan en momentos diferentes del ciclo vital del virus VIH.

1.4. Diseño y optimización del zimógeno ONCFLG de la Onconasa®

En el diseño del zimógeno, se tuvieron en cuenta dos características de la ONC relacionadas con sus extremos N- y C- terminales, los cuales pueden influir en la actividad y sobre todo en la estabilidad y citotoxicidad de la misma. Primero, se conoce que en el extremo N- terminal, el primer residuo es una glutamina (Gln1) que espontáneamente se cicla para formar un residuo de ácido piroglutámico (Pyr). La estructura cristalográfica de la ONC muestra como este residuo se pliega hacia la hélice $\alpha 1$ y forma un enlace de hidrógeno con el grupo CO de la Val96 que se encuentra cercano en la lámina β del extremo C-terminal, figura 5.5 (A). Pyr1 también forma un enlace de hidrógeno con el grupo ϵ -amino de la cadena lateral de la Lys9, el cual interactúa con el sustrato durante la catálisis (Mosimann et al. 1994; Lee & Raines 2003). Por su naturaleza cíclica, este Pyr no puede conservarse en el diseño del zimógeno. Por ello fue reemplazado por un residuo de Ser, ya que se ha descrito (Liao et al. 2003) que el mutante de ONC que lleva esta modificación es el que presenta la mayor eficiencia catalítica y citotóxica. Segundo, la ONC a diferencia de la RNasa A, presenta un enlace disulfuro sinapomórfico C87/C104 que une la Cys104 C-terminal a la lámina β central y la aísla del extremo N-terminal. Ello implica que no se pueden unir los extremos N- y C- directamente (Mosimann et al. 1994) sin modificar su plegamiento. Un análisis de la estructura cristalográfica de la ONC (código pdb 1ONC) revela que la distancia entre los extremos C- y N- terminales es de 18 Å. Esta distancia podría ser cubierta por un polipéptido de 5 residuos si estuviesen en una conformación totalmente extendida. Sin embargo, los estudios de modelado, figura 5.5 (B), realizados por el grupo de Girona, sugieren que serían necesarias secuencias de 14 o más aminoácidos para conectar los extremos nativos y bloquear el sitio activo, sin perturbar la estructura global. La predicción de estructura llevada a cabo con este fin utilizando el servidor SWISS-MODEL, conduce a un modelo en el que el segmento de unión adquiere una estructura helicoidal.

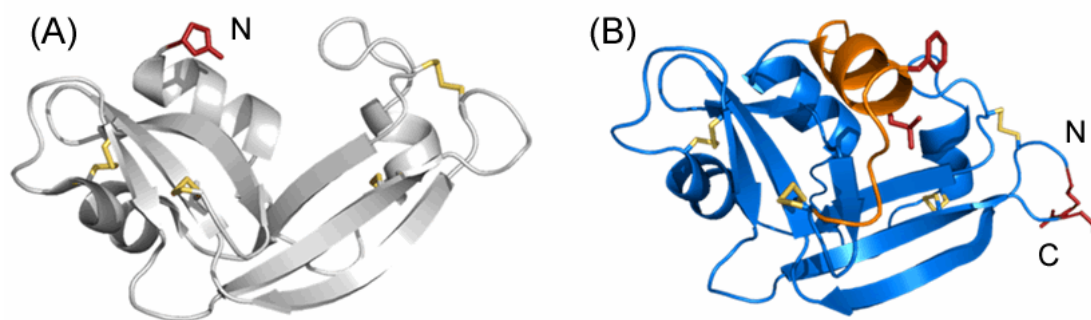


Figura 5.5. Representación de la ONC y del modelo predicho del zimógeno con la secuencia de unión. (A) Representación en forma de cinta de la Onconasa® (código de acceso, 1ONC). El extremo N-terminal está representado en rojo y los cuatro enlaces disulfuro en amarillo. (B) Modelo estructural del zimógeno ONCFLG generado usando el servidor de modelización SWISS-MODEL y minimizado energéticamente con GROMOS 96, implementación de Swiss-PdbViewer (Fyfe 1983). La secuencia insertada está representada en naranja. Se indican los nuevos extremos N- y C- terminal. Las cadenas laterales de los residuos que constituyen el sitio de corte se muestran en rojo y los enlaces disulfuro en amarillo. Ambas figuras han sido dibujadas utilizando el programa PyMOL (<http://pymol.sourceforge.net/>) y han sido publicadas (Callís et al. 2013).

La proteasa seleccionada en este trabajo, la proteasa específica del VIH-1, es una aspartil proteasa que cataliza la escisión de los precursores de la poliproteína Pr55^{gag} y Pr160^{gag-pol} que codifican las proteínas estructurales y enzimas del virus. Esta actividad proteolítica es absolutamente indispensable para la producción de los viriones infecciosos maduros (Debouck 1992). Después de varias optimizaciones experimentales por el grupo del Dr. Ribó, la secuencia escogida para ser insertada y unir los extremos terminales de ONC fue: GSGGSGIF*LETSLSAG. En ella, la secuencia de reconocimiento para la proteasa el VIH-1 aparece subrayada y el asterisco indica el enlace peptídico hidrolizado por la misma. De ahora en adelante nos referiremos a esta construcción como zimógeno ONCFLG, figura 5.6.

Los residuos Arg73 y Ser72 de la ONC se escogieron como nuevos extremos N- y C-terminales basándose en la elevada accesibilidad de los mismos en el bucle que conecta las láminas β_4 - β_5 de la estructura. En este bucle se pueden realizar sustituciones sin que se altere la estructura de la ONC. Por ejemplo, se ha sustituido con éxito la Ser72 por Cys, y su buena estabilidad le permite ser usada como punto de anclaje de un marcador de fluorescencia y poder seguir de esta manera la internalización celular de la proteína (Rodríguez et al. 2007).

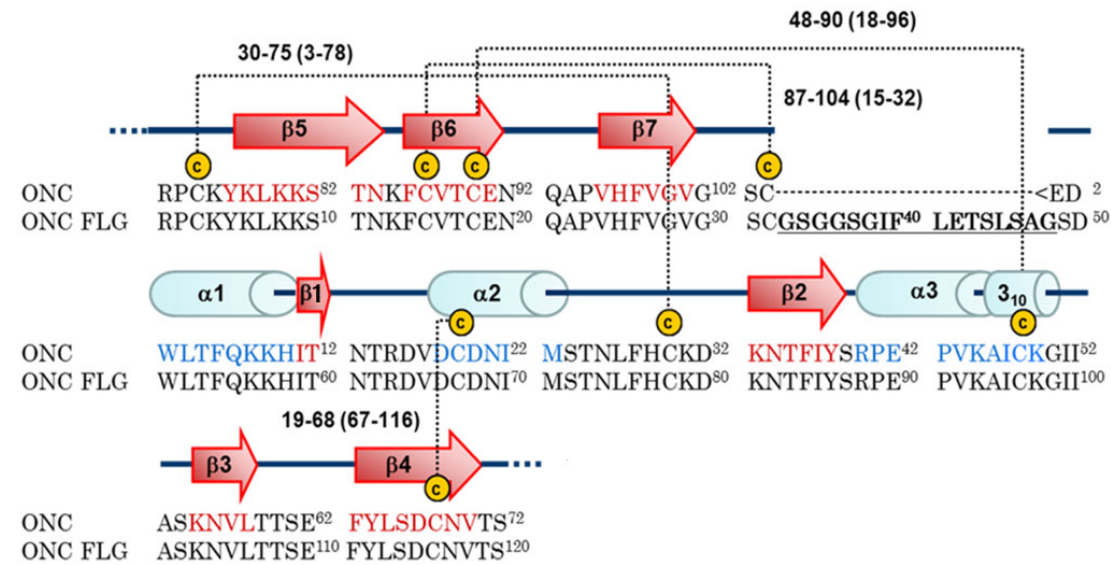


Figura 5.6. Alineamiento de la secuencia de aminoácidos de la ONC y ONCFLG representando la estructura secundaria de la ONC. Los residuos que forman la secuencia de unión entre Ser1 y Cys104 están representados en negrita y subrayados. La figura ha sido tomada del artículo (Callís et al. 2013).

1.5. Objetivo

Los fines terapéuticos que se pretenden explorar con la aproximación que se ha presentado, se basan en que el zimógeno se internalice adecuadamente y que una vez en el interior de la célula infectada pueda llevar a cabo su acción y consecuentemente lograr la muerte celular selectiva. Por ello es interesante poder describir el proceso de interacción con la membrana como paso fundamental de la internalización, estudiar a nivel atómico los cambios que se producen en el proceso, las interacciones que tienen lugar con sus componentes y procesos asociativos si los hubiera, y finalmente conocer cómo se recupera o como se mantiene la estructura activa del zimógeno dentro de la célula.

Como un paso inicial en este objetivo último que podría desarrollarse a más largo plazo, en este capítulo nos hemos propuesto caracterizar a nivel atómico un zimógeno en sus formas hidrolizada y sin hidrolizar como punto de partida de un diseño mejorado que presente buenas posibilidades terapéuticas.

Dadas las especiales características de la Onconasa® y considerando su potencial, en el Grupo de Ingeniería de Proteínas de la Facultad de Biología de la Universitat de Girona con el que colaboramos, se han diseñado, producido y caracterizado varias permutaciones circulares de esta proteína que pueden ser selectivamente activadas por la proteasa VIH-1.

De las variantes producidas por el grupo del Prof. Ribó, el zimógeno de Onconasa® ONCFLG, presentó en general las mejores características de internalización (lo que es importante para nuestros estudios futuros), activación y actividad, mostrándose como un buen candidato para nuestros fines. Por estos motivos se escogió esta construcción para realizar un estudio encaminado a comprender las bases estructurales y dinámicas de la actividad catalítica y la alta estabilidad de esta peculiar enzima en sus formas completa e hidrolizada. Utilizando la RMN como herramienta, se ha llevado a cabo un análisis completo con muestras isotópicamente marcadas y aplicando métodos heteronucleares 2D y 3D RMN. Además de la determinación de la estructura tridimensional, se han realizado estudios de la dinámica interna y la estabilidad conformacional por residuo.

Los cambios estructurales y dinámicos resultantes de la activación del zimógeno ONCFLG por la proteasa del VIH-1 deben permitir describir la relación estructura-actividad, así como comprender los motivos estructurales responsables de su alta estabilidad conformacional, los cuales pueden ser importantes desde el punto de vista de la internalización celular y su posible acción terapéutica.

Los resultados de este capítulo se han publicado en colaboración con el grupo del Prof. Marc Ribó del Laboratori d'Enginyeria de Proteïnes, del Departament de Biologia, Facultat de Ciències Universitat de Girona: Callís M., Serrano S., Benito A., Laurents D.V., Vilanova M., Bruix M. and Ribó M. *"Towards Tricking a Pathogen's Protease into Fighting Infection: The 3D Structure of a Stable Circularly Permuted Onconase Variant Cleaved by HIV-1 Protease"* PLoS One 8 **(2013)** (8) e54568. Algunas de las imágenes incluidas en este texto pertenecen a dicha publicación, como se indica en los correspondientes pies de figura. La asignación de los espectros de RMN del zimógeno ONCFLG ha sido también publicada: Serrano S., Callis M., Vilanova M., Benito A., Laurents D.V., Ribó M., Bruix M. *"(1)H, (13)C and (15)N resonance assignments of the Onconase FL-G zymogen"* Biomol. NMR Assign. **(2013)** 7,13-15.

2. MATERIALES Y MÉTODOS

2.1. Preparación de la muestra para RMN del zimógeno ONCFLG

Para su estudio por RMN, el zimógeno recombinante ONCFLG fue producido y purificado a partir de cultivos de células de *Escherichia coli* BL21(DE3) siguiendo un proceso que había sido descrito previamente (Ribo et al. 2004). Se prepararon tres tipos de muestras, (I) sin marcaje isotópico, (II) con ^{15}N únicamente, y (III) con doble etiquetado uniforme de ^{15}N y ^{13}C .

Las muestras de RMN se prepararon disolviendo 18.5 mg de proteína, cantidad necesaria para tener una concentración de 2.8 mM del zimógeno ONCFLG, en 0.45 mL H_2O y 0.05 mL de D_2O . Las muestras contenían, además, 50 μM DSS (2,2-dimetil-2-silapentano-5-sulfonato de sodio), 3 mM NaN_3 , 100 mM KCl, 1 mM HAc, 1 mM K_2HPO_4 .

El pH de la muestra se ajustó a 5.20 mediante la adición de las alícuotas necesarias de DCl o NaOD. El valor de pH se midió con un microelectrodo Hamilton de 3 mm de diámetro, y no se aplicó ninguna corrección por efectos isotópicos.

2.2. Hidrólisis del zimógeno ONCFLG

Para producir el corte en el enlace específico se ha utilizado la proteasa de VIH recombinante producida en el laboratorio del Prof. Marc Ribó de la Universitat de Girona. La eficiencia catalítica de esta proteasa es comparable a la de la proteasa comercial.

La activación del zimógeno por la proteasa del VIH, fue llevada a cabo realizando ensayos con distintas relaciones proteasa/zimógeno (relaciones de 1/5 a 1/100), en un tampón 100 mM compuesto por acetato sódico/ácido acético a pH 4.7, NaCl 300 mM y EDTA 4 mM.

Empleando la relación molar más baja proteasa/zimógeno (relación 1/5), se tomaron alícuotas a distintos tiempos de incubación, las cuales fueron mezcladas con SDS para parar la reacción de hidrólisis y se almacenaron a $-20\text{ }^\circ\text{C}$ hasta que se procedió a su análisis por electroforesis con gel de poliacrilamida y SDS, SDS-PAGE. La reacción de corte del zimógeno fue cuantificada utilizando Imaging System FluorChem® SP (Alpha Innotech, San Leandro, CA, USA).

Experimento	Muestra	Nº barridos	Anchura espectral ppm/ppm (ptos/ptos)	Correlaciones observadas
2D		NS	t₂/t₁	
¹ H- ¹⁵ N HSQC	¹⁵ N	8	12/21 (2Kx128)	¹ H _{N(i)} , ¹⁵ N _{H(i)}
¹ H- ¹³ C HSQC	¹⁵ N, ¹³ C	4	12/80 (2Kx256)	¹ HC _(i) , ¹³ C _{alif.(i)}
¹ H- ¹³ C HSQC (aromáticos)	¹⁵ N, ¹³ C	8	16/30 (2Kx256)	¹ HC _(i) , ¹³ C _{arom.(i)}
¹ H- ¹ H NOESY		64	12/12 (2Kx512)	¹ H, ¹ H (dist.)
(HB)CBHE (aromáticos)	¹⁵ N, ¹³ C	80	12/24 (2Kx128)	¹ HC _(i) , ¹³ C _{aro(i)}
(HB)CBHD (aromáticos)	¹⁵ N, ¹³ C	80	12/24 (2Kx128)	¹ HC _(i) , ¹³ C _{aro(i)}
3D		NS	t₃/t₂/t₁	
HNCA	¹⁵ N, ¹³ C	8	12/21/35 (2Kx26x100)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ C _{α(i,i-1)}
HNCO	¹⁵ N, ¹³ C	4	12/21/24 (2Kx26x80)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ CO _(i-1)
HN(CO)CA	¹⁵ N, ¹³ C	8	12/21/35 (2Kx26x100)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ C _{α(i-1)}
HN(CA)CO	¹⁵ N, ¹³ C	8	12/21/24 (2Kx26x80)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ CO _(i,i-1)
CBCANH	¹⁵ N, ¹³ C	16	12/21/75 (2Kx26x128)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ C _{α(i,i-1)} / ¹³ C _{β(i,i-1)}
CBCA(CO)NH	¹⁵ N, ¹³ C	8	12/21/70 (2Kx26x128)	¹ H _{N(i)} , ¹⁵ N _{H(i)} , ¹³ C _{α(i-1)} / ¹³ C _{β(i-1)}
HC(C)H-TOCSY	¹⁵ N, ¹³ C	8	12/75/12 (2Kx64x128)	¹ H _(i) , ¹³ C _(i) , ¹ H _(i)
(H)CCH-TOCSY	¹⁵ N, ¹³ C	8	12/75/75 (2Kx75x75)	¹³ C _(i) , ¹³ C _(i) , ¹ H _(i)
¹ H- ¹⁵ N HSQC NOESY	¹⁵ N, ¹³ C	8	12/21/12 (2Kx64x160)	¹ H, ¹⁵ N, ¹ H (distancias)
¹ H- ¹³ C HSQC NOESY (alifáticos)	¹⁵ N, ¹³ C	16	16/80/16 (2Kx64x128)	¹ H, ¹³ C, ¹ H (distancias)
¹ H- ¹³ C HSQC NOESY (aromáticos)	¹⁵ N, ¹³ C	8	12/40/12 (2Kx64x160)	¹ H, ¹³ C, ¹ H (distancias)

Tabla 5.1. Experimentos adquiridos para el estudio por RMN del zimógeno ONCFLG. En la segunda columna se indica el etiquetado de la muestra empleada en la adquisición del experimento.

2.3. Adquisición de los espectros de RMN para cálculo de estructura

Los espectros necesarios para la asignación espectral y el cálculo de la estructura tridimensional del zimógeno ONCFLG se adquirieron a 35 °C en un espectrómetro Bruker Avance 800 MHz con un campo magnético de 18.8 Teslas que opera a 800.13 MHz para el protón, equipado con una criosonda TXI con gradientes en el eje z. La temperatura de la sonda de RMN se calibró utilizando una muestra de metanol (Duerst & Verbach 1965). Los desplazamientos químicos de protón fueron referenciados internamente respecto a la señal del DSS. Los desplazamientos heteronucleares se referenciaron de manera indirecta (Wishart et al. 1995a). La supresión del agua se llevó a cabo por presaturación selectiva de la misma o incluyendo un módulo WATERGATE (Piotto et al. 1992) al comienzo de la secuencia de pulsos, antes de iniciar la adquisición.

La tabla 5.1, en la página anterior, recoge los distintos experimentos de RMN que han sido adquiridos para llevar a cabo el estudio estructural. Las secuencias de pulsos utilizadas son las estándares de la aproximación multidimensional heteronuclear (Whitehead et al. 1997).

2.4. Procesado y análisis de los espectros de RMN

Todos los espectros de RMN fueron procesados usando el programa Topspin 2.1 (Bruker Biospin, Karlsruhe, Alemania). Durante el procesado se incrementó el número de puntos según conveniencia y se utilizaron funciones seno o seno cuadrado para la apodización. Finalmente se realizó la transformada de Fourier para cada experimento y se corrigió la fase y la línea base.

Tanto para la visualización como para el análisis de los espectros se utilizó el programa SPARKY (Goddard 2005).

2.5. Asignación espectral

Experimento	Información obtenida
^1H - ^{15}N HSQC	Asignación del esqueleto y átomos de la cadena principal $^{15}\text{N}_\text{H}$, $^1\text{H}_\text{N}$, $^{13}\text{C}_\alpha$, C' y $^{13}\text{C}_\beta$
HNCA	
HN(CA)CO	
HN(CO)CA	
CBCANH	
CBCA(CO)NH	
^1H - ^{13}C HSQC	Asignación de resto de ^1H y ^{13}C de las cadenas laterales
NHCO	
^1H - ^{13}C HSQC	
^1H - ^{13}C HSQC (aromat.)	
HC(C)H-TOCSY	
(H)CCH-TOCSY	
(HB)CBHD (aromat.)	Asignación de átomos de ^1H y ^{13}C aromáticos
(HB)CBHE (aromat.)	
^1H - ^{13}C HSQC (aromat.)	
^1H - ^1H NOESY	Restricciones de distancia
^1H - ^{15}N HSQC NOESY	
^1H - ^{13}C HSQC NOESY	
^1H - ^{13}C HSQC NOESY (aromat.)	

Tabla 5.2. Relación entre los experimentos de RMN adquiridos y la información obtenida de los mismos para la determinación estructural del zimógeno ONCFLG.

La asignación secuencial de proteínas se aborda a través de diferentes estrategias, una está basada en los experimentos de correlación homonuclear ^1H - ^1H (Wüthrich 1986), que es la utilizada en los capítulos III y IV, y otra basada en correlaciones heteronucleares de proteínas enriquecidas en ^{13}C y ^{15}N (Souza et al. 2004), que es la empleada para ONCFLG.

En este caso, con la muestra marcada en ^{13}C y ^{15}N la asignación de la proteína se comenzó por los átomos de la cadena principal, $^1\text{H}_\text{N}$, $^{13}\text{C}_\alpha$, $^{13}\text{C}'$ y $^{15}\text{N}_\text{H}$ y $^{13}\text{C}_\beta$, mediante el análisis de una serie de experimentos como son el CBCANH, CBCA(CO)NH, HN(CO)CA, HNCA, HN(CA)CO, HNCO. Una vez hecho esto, se realizó la asignación de los $^1\text{H}_\alpha$, $^1\text{H}_\beta$ y el resto de ^1H y ^{13}C de las cadenas laterales alifáticas de los aminoácidos empleando los experimentos 3D HC(C)H-TOCSY y (H)CCH-TOCSY. Para la asignación de los carbonos de los anillos aromáticos se utilizó un espectro 2D ^1H - ^{13}C HSQC realizado específicamente en la región de los mismos y para los protones de las cadenas laterales aromáticas se emplearon

los experimentos 2D (HB)CBHD y (HB)CBHE, a partir de las correlaciones de los protones aromáticos con el C_β , tabla 5.2.

2.6. Cálculo de estructuras

La determinación de la estructura de proteínas en solución mediante RMN se realiza tras un primer paso de asignación secuencial de las señales y la determinación de la estructura secundaria; y un segundo paso, en el que los parámetros de RMN (δ , J, NOE y RDC) se convierten en restricciones geométricas que deben cumplir los átomos de la proteína. En este caso los cálculos estructurales han sido realizados utilizando los NOEs del experimento 2D 1H - 1H NOESY con un tiempo de mezcla de 80 ms, y de los experimentos 3D 1H - ^{15}N HSQC NOESY, NOESY 1H - ^{13}C HSQC (alifáticos) y 1H - ^{13}C -HSQC NOESY (de la región de los aromáticos), todos ellos con un tiempo de mezcla de 50 ms.

Además de las restricciones de distancia obtenidas a partir de las señales de los espectros tipo NOESY, se incluyeron restricciones angulares calculadas mediante el programa TALOS (Cornilescu et al. 1999). TALOS, realiza una predicción de los valores de los ángulos diedros del esqueleto, Φ y ψ , en función de los desplazamientos químicos de $^1H_\alpha$, $^{13}C_\alpha$, $^{13}C_\beta$, $^{13}C'$ y $^{15}N_H$ a partir de una base de datos que emplea desplazamientos químicos y residuos homólogos.

El cálculo consiste en encontrar el conjunto de estructuras cuyas coordenadas atómicas cumplen las restricciones experimentales, de distancia, angulares, y de orientación (en el caso de que hayan sido determinadas) proporcionadas por los parámetros de RMN.

Los cálculos para determinar la estructura del zimógeno ONCFLG han sido realizados utilizando el programa CYANA 2.1 (Güntert 2004). El programa CYANA, se ha utilizado con el protocolo de asignación semiautomática de señales NOE, conjuntamente con la inclusión de asignaciones manuales de los mismos. En el cálculo se generaron 100 confórmeros, y se seleccionaron los 20 mejores, es decir, los que poseen menor valor de función blanco, como los representativos de la estructura del zimógeno en disolución.

2.7. Análisis de estructuras

El programa CYANA (Güntert 2004) no sólo nos permite realizar los cálculos sino que nos proporciona un serie de herramientas para el análisis de las estructuras: número y magnitud de las violaciones, el valor de la desviación cuadrática media entre confórmeros de la familia (RMSD), etc, que se han utilizado en el análisis. Estas violaciones vienen dadas por restricciones, ya sean de distancia o angulares, que no se cumplen adecuadamente en la

familia de estructuras obtenida. En las estructuras resultantes, el número y magnitud de las violaciones de las restricciones, debe ser mínimo. La convergencia de la familia de estructuras calculadas, se mide por el valor de la desviación cuadrática media (RMSD) entre pares. Valores bajos de RMSD indican buena convergencia y valores altos indican variabilidad conformacional entre las mismas.

La representación y el análisis de las estructuras se realizó también utilizando los programas MOLMOL (Koradi et al. 1996) y VMD (Humphrey et al. 1996).

2.8. Minimización

La calidad de las estructuras, que han sido obtenidas mediante restricciones geométricas, puede ser mejorada incluyendo consideraciones energéticas, por ello, una vez obtenidas las estructuras con CYANA, procedemos a una minimización de energía.

Mediante la aplicación de algoritmos de minimización de energía se pretende la búsqueda de una configuración del sistema de energía mínima. En nuestro caso particular, esta minimización energética, se llevó a cabo con el programa AMBER (Case 2006).

Los cálculos con AMBER se realizaron en condiciones de vacío y con disolvente explícito, de manera que los valores de energía obtenidos se utilizan para describir las diferentes contribuciones a la energía global del sistema.

2.9. Validación y depósito

Programas bien conocidos como AQUA y PROCHECK (Laskowski et al. 1996) tienen por objeto realizar un gran número de comprobaciones de validación en un conjunto de estructuras de proteínas obtenidas por RMN. En este trabajo, se ha empleado PROCHECK, para obtener la calidad de la geometría obtenida y la adecuación a las restricciones experimentales del conjunto de estructuras obtenidas por CYANA.

Una vez realizada esta última validación las estructuras obtenidas han sido depositadas en la base de datos del PDB (Protein Data Bank, <http://www.rcsb.org/pdb/>) con código de acceso 2LT5.

2.10. Interacción de ONCFLG con el análogo de sustrato d[UGG]₃

El oligonucleótido d[UGG]₃, adquirido en Thermo-Fischer Scientific (Ulm, Alemania), representa un buen sustrato no hidrolizable para la Onconasa®.

La muestra de RMN se preparó con concentraciones de $d[UGG]_3$ y ^{15}N -ONCFLG de 0.63 y 0.32 mM respectivamente, en 0.45 mL de H_2O y 0.05 mL de D_2O . La muestra, contenía también, 0.1 M de KCl, 50 μM de DSS y 0.6 de NaN_3 . El pH de la misma fue ajustado a 5.2 con una mezcla de ácido acético y acetato sódico deuterados ($NaAc(d_3)/HAc(d_3)$), a 6.35 y 7.25 con 1mM de K_2HPO_4 y 1mM de borato sódico respectivamente.

Se registraron experimentos 1D 1H y 2D 1H - ^{15}N HSQC, centrados en la región del esqueleto y los grupos amida de las cadenas laterales de Asn y Gln o, alternativamente, en el imidazol de las His, 2D 1H - 1H TOCSY, 2D 1H - 1H NOESY, 3D 1H - ^{15}N HSQC TOCSY y 3D 1H - ^{15}N HSQC NOESY. Finalmente, se adquirió una serie similar en una muestra de ONCFLG a pH 7.28, en el mismo tampón, pero sin $d[UGG]_3$.

En un intento por favorecer la interacción entre el ^{15}N -ONCFLG y $d[UGG]_3$, la muestra fue desalada mediante diálisis. La muestra resultó insoluble a pH 7 y los experimentos sin sal se adquirieron a pH 7.9, donde resultó moderadamente soluble. Finalmente, se subió el pH a 9.0, por adición de NaOD. En este punto la muestra era completamente soluble y se adquirió la serie final de experimentos del posible complejo: 1D de 1H y 2D 1H - 1H NOESY y 1H - ^{15}N HSQC NOESY.

2.11. Construcción del modelo estructural del zimógeno hidrolizado

En primer lugar se compararon los espectros 2D-RMN de ONCFLG en la forma intacta e hidrolizada y se comprobó que los espectros no presentaban variaciones importantes en la posición de las señales, lo que indica que los cambios producidos por la ruptura del enlace entre Phe40 y Leu41 (diana de corte) son locales y no afectan al plegamiento global. A partir de esta evidencia, el cálculo se llevó a cabo con el programa CYANA 2.1 (Güntert 2004) a partir de la secuencia polipeptídica inicial y simulando la hidrólisis del enlace 40-41 con las herramientas de CYANA adecuadas. Las restricciones empleadas en el cálculo incluyen las utilizadas con el zimógeno no hidrolizado, excluyendo las de la zona adyacente a la hidrólisis que van desde la Gly33 hasta Gly48.

2.12. Adquisición de los espectros de RMN para la dinámica

Para caracterizar la dinámica del esqueleto de la proteína, en sus formas hidrolizada y no hidrolizada, se ha evaluado el NOE heteronuclear 1H - ^{15}N . Estas medidas se realizaron usando muestras de la ONCFLG marcada en ^{15}N a pH 5.2 y una temperatura de 35 °C.

Los espectros con y sin saturación de protón fueron adquiridos simultáneamente de forma intercalada y durante el procesamiento se separaron en espectros independientes. El tiempo de espera entre barridos que se utilizó fue de 5.0 segundos.

Los valores de NOE heteronuclear se obtuvieron del cociente entre las intensidades de ambos espectros ($I_{\text{sat}}/I_{\text{nosat}}$). Las intensidades y volúmenes de los picos de los espectros bidimensionales se evaluaron usando herramientas del programa SPARKY (Goddard 2005).

2.13. Adquisición de los espectros de RMN para el estudio de estabilidad conformacional

Las velocidades de intercambio H/D de los protones amida nos proporcionan información muy útil acerca de la participación de dichos protones en enlaces de hidrógeno y, además, nos permiten determinar la estabilidad conformacional de la proteína, a nivel de residuo.

El intercambio protón/deuterio, de los protones amídicos con el solvente deuterado, tanto del zimógeno ONCFLG hidrolizado como sin hidrolizar, se inició disolviendo el liofilizado en D₂O. Los experimentos se realizaron a una temperatura de 30 °C y a un pH de 6.0, para el zimógeno sin hidrolizar, y 5.2, para el zimógeno hidrolizado. Estos valores de pH se escogieron porque resultaban los adecuados para que el intercambio fuese lo bastante rápido para poder medir los grupos con procesos más lentos y para que el intercambio ocurriese dentro del régimen EX2.

Se realizaron series de ¹H-¹⁵N HSQC consecutivos. El tiempo entre cada experimento fue de 30 min. Los experimentos se adquirieron con 2048 puntos en t_2 y con 128 en t_1 , respectivamente y con 8 barridos. Cada experimento individual duró aproximadamente 15 minutos.

Las constantes de intercambio de hidrógeno, k_{ex} , fueron determinadas a partir de la integración del volumen de los picos de ¹H-¹⁵N amídicos utilizando una rutina integrada en el programa SPARKY (Goddard et al. 2005). El programa OriginLab, Origin 6 se empleó para el ajuste con una única función de decaimiento de las intensidades en función del tiempo de experimento.

Los valores intrínsecos de intercambio, k_{rc} , se calcularon usando los parámetros reportados en la literatura (Bai et al. 1993), corregidos en función del pH y la temperatura. Usando las ecuaciones descritas en este último trabajo, se determinaron los factores de protección, P.

Cuando es el régimen EX2 el que gobierna el intercambio (condiciones estabilizadoras, baja temperatura, bajo pH) la estabilidad conformacional, ΔG_{HX} , para cada grupo amida puede ser calculada a partir del factor de protección mediante la expresión:

$$\Delta G_{HX} = -RT \ln(K_{ex}/k_{rc}) = -RT \ln(1/P) = -RT \ln K_{op}$$

dónde K_{op} es la constante de equilibrio local o global para la reacción de apertura que permite el intercambio.

3. RESULTADOS

3.1. Caracterización preliminar de la proteína

Para comprobar el estado de plegamiento y seleccionar unas condiciones adecuadas para el estudio estructural, el primer paso a realizar fue la adquisición de un experimento monodimensional de RMN.

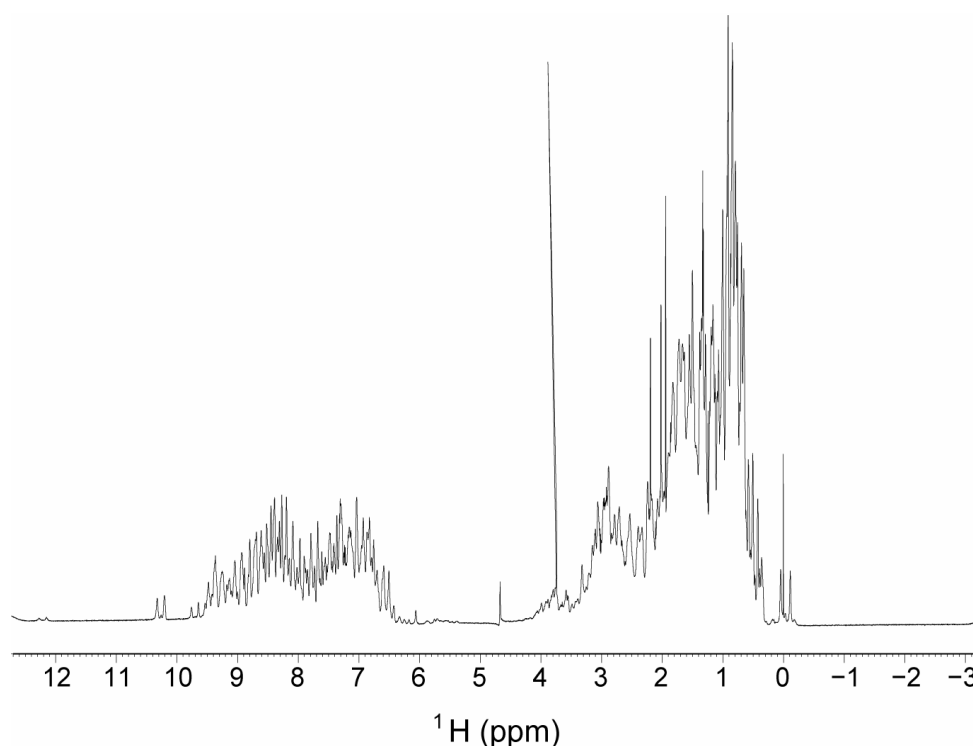


Figura 5.7. Espectro 1D del zimógeno ONCFLG marcado en ^{15}N , concentración 2.8 mM en 0.45 mL H_2O y 0.05 mL de D_2O a pH 5.2 y 35 °C.

Tal y como puede verse en la figura 5.7, el espectro de RMN de ^1H presenta una gran dispersión de señales, principalmente en la región correspondiente a los protones H_N amida, que se encuentra comprendida entre 10.5 y 5.5 ppm. También es evidente esta dispersión en las zonas más apantalladas del espectro, a desplazamientos químicos inferiores a 0.5 ppm. Por tanto podemos concluir, que en estas condiciones experimentales (concentración 2.8 mM, pH 5.2 y temperatura de 35 °C) la proteína se encuentra bien plegada por lo que son adecuadas para el estudio y adquisición de los experimentos de RMN para el zimógeno ONCFLG no hidrolizado.

3.2. Asignación espectral

Las asignaciones de los desplazamientos químicos de ^1H , ^{15}N y ^{13}C de la proteína se han realizado utilizando los espectros que se resumen en la tabla 5.2 del apartado de Materiales y Métodos.

La secuencia del zimógeno ONCFLG está compuesta por un total de 120 aminoácidos, 104 corresponden a los de la Onconasa® silvestre y 16 a la secuencia insertada que sirve de unión de los extremos N- y C- terminal nativos. En la secuencia de la proteína encontramos 4 Pro en la posiciones Pro2, Pro23, Pro89 y Pro91. Las cisteínas están formando cuatro enlaces disulfuro: Cys3-Cys78, Cys15-Cys32, Cys18-Cys96 y Cys67-Cys118.

Para realizar la asignación completa, primero se han asignado los átomos de la cadena principal de la proteína $^1\text{H}_\text{N}$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, ^{13}CO y $^{15}\text{N}_\text{H}$, figura 5.8, y a continuación $^1\text{H}_\alpha$ y resto de los núcleos de las cadenas laterales.

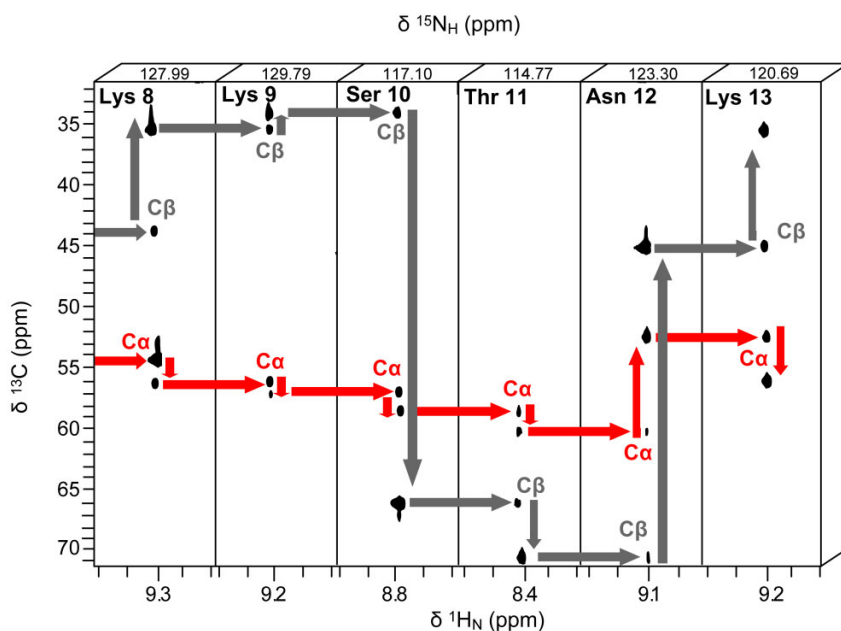
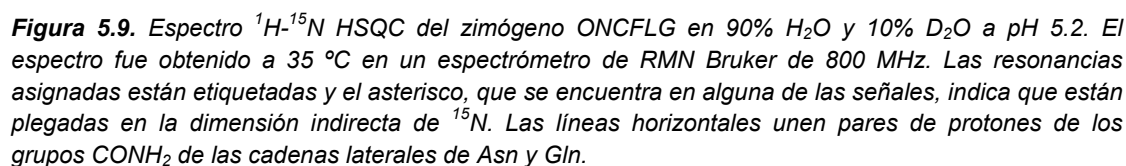


Figura 5.8. Asignación secuencial del fragmento que incluye a los residuos Lys8, Lys9, Ser10, Thr11, Asn12 y Lys13 empleando el espectro 3D CBCANH. En este tipo de experimento, en un mismo plano de ^{15}N se observan las señales correspondientes a los C_α y C_β de un residuo (i) y del anterior (i-1). La flechas rojas muestran la asignación secuencial de los C_α y las grises la de los C_β .

Como se observa en el espectro bidimensional ^1H - ^{15}N HSQC de la figura 5.9, fue posible identificar todas las señales de las correlaciones de ^1H - ^{15}N del esqueleto excepto la que corresponde a la Phe26. En estudios anteriores en los que se describen las asignaciones de RMN de variantes de la Onconasa® (Gorbatyuk et al. 2004; Schulenburg et al. 2009),

Se identificaron ocho pares de correlaciones de cadenas laterales ^1H - ^{15}N , de los cuales, seis corresponden a los residuos de asparaginas (Asn50, Asn64, Asn66, Asn68, Asn80 y Asn115) y dos a los residuos de glutamina (Gln21 y Gln55) de la secuencia.



Una vez determinados los desplazamientos químicos de los átomos de la cadena principal de ONCFLG, se completó la asignación de los protones de las cadenas laterales empleando los espectros 3D HC(C)H-TOCSY y (H)CCH-TOCSY para cadenas alifáticas y experimentos 2D ^1H - ^{13}C HSQC y 2D (HB)CBHD y (HB)CBHE, para los carbonos y protones aromáticos, respectivamente.

Nuestro análisis ha revelado algunos valores inusuales de desplazamientos químicos cuando se comparan con los valores estándares (Wishart et al. 1995b). Este es el caso de un protón amida δ de Asn82 (3.64 ppm), un protón amida ϵ de Gln55 (4.67 ppm), H_γ de Lys93 (-0.09 ppm), Me_γ de Val65 (-0.02 ppm) y C_α de Glu90 (63.59 ppm). El caso del Glu90 es particular ya que, al encontrarse entre dos Pro, su asignación resultó difícil.

La asignación de las resonancias del esqueleto se ha completado en un 97.8% mientras que la asignación de las resonancias de las cadenas laterales lo ha sido en un 94%. Existe todavía un pequeño grupo de resonancias en las que no es posible la asignación inambigua, concretamente grupos NH de las cadenas laterales de algunos sistemas de espín de Lys y Arg.

La asignación de los desplazamientos químicos para el esqueleto y las cadenas laterales del zimógeno ONCFLG (Serrano et al. 2012) ha sido depositada en la base de datos del BioMagResBank (<http://www.bmrb.wisc.edu>), BMRB, con el código de acceso 17973.

3.3. Estructura del zimógeno ONCFLG

3.3.1. Estructura secundaria

La estructura secundaria de péptidos y proteínas puede identificarse utilizando los desplazamientos conformacionales de los $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ y las intensidades de los NOEs secuenciales y de medio alcance, como se ha mencionado en el apartado 2.7 del capítulo III.

La figura 5.10 muestra los desplazamientos químicos conformacionales de los núcleos de $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$. De acuerdo con esta información, se observan tres regiones helicoidales: Leu52-His58, Ile70-Ser72 y Pro89-Ala94 (en la numeración del zimógeno); con desplazamientos químicos conformacionales positivos (valores de $\Delta\delta$ entre +1 y +7 ppm) para los $^{13}\text{C}_\alpha$ y valores negativos ($\Delta\delta$ entre -0.5 y -4.5 ppm) para los $^{13}\text{C}_\beta$. Hay siete regiones en lámina beta: Tyr5-Asn12, Phe14-Cys18, Pro23-Val29, Ile59-Thr60, Lys81-Tyr86, Lys103-Thr107 y Phe111-Val118 (en la numeración del zimógeno); los cuales presentan valores positivos ($\Delta\delta$ entre +1 y +8 ppm) para los $^{13}\text{C}_\beta$ y valores negativos ($\Delta\delta$ entre -0.5 y +4 ppm) para los $^{13}\text{C}_\alpha$.

A pesar de haber sido diseñada como helicoidal, la secuencia de 16 aminoácidos insertada entre los extremos nativos N- y C- terminal (del residuo 33 al 48 en la numeración del zimógeno) no adopta ninguna estructura secundaria, como muestran los bajos valores de desplazamiento conformacional (valores de $\Delta\delta$ de C_α entre -0.28 y +0.28 ppm).

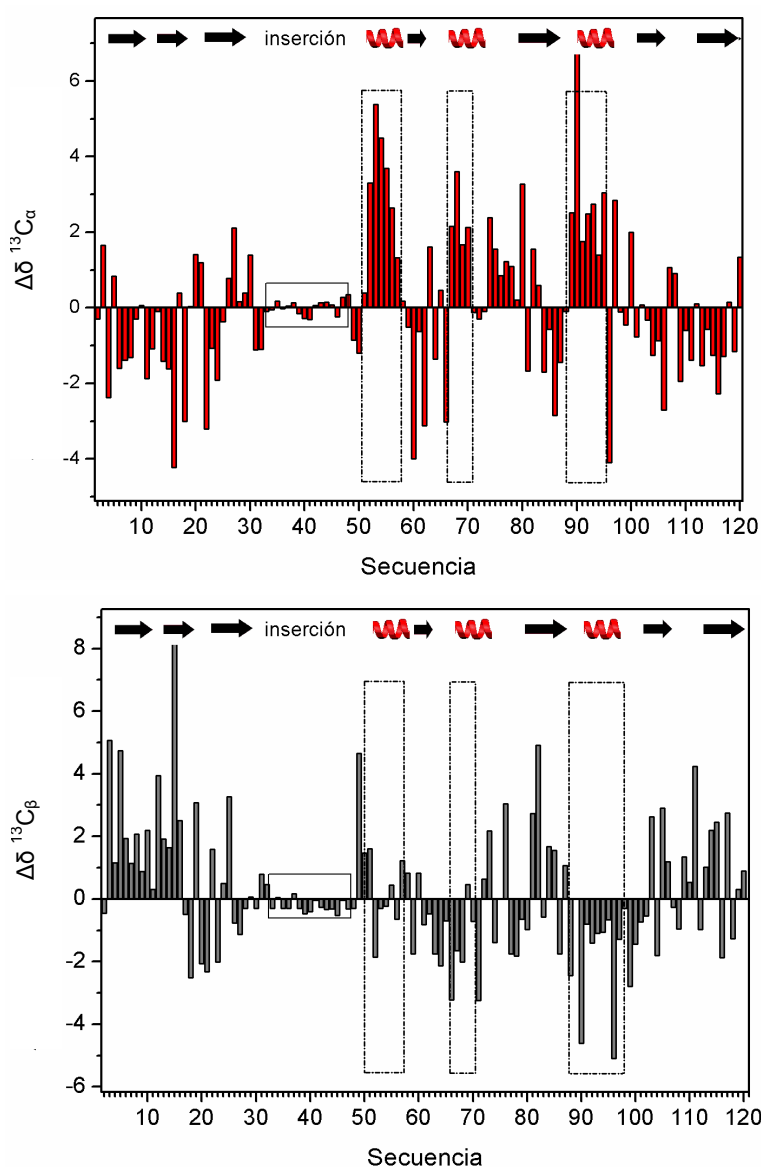


Figura 5.10. Desplazamientos químicos conformacionales para los C_α y C_β del zimógeno en función de la secuencia de la proteína. Los valores de $\Delta\delta$ son la diferencia entre los valores observados y los valores en péptidos sin estructura ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm). Los valores $\delta_{\text{random coil}}^{C_\alpha}$ y $\delta_{\text{random coil}}^{C_\beta}$ de referencia se tomaron de Wishart et al. (Wishart et al. 1995a). En la parte superior está representada la estructura secundaria: las hebras β como flechas negras y las hélices α como espirales rojas. El rectángulo que abarca del residuo 33 al 48 indica la secuencia insertada y los rectángulos punteados las zonas con estructura de hélice.

De los experimentos 3D ^1H - ^{15}N HSQC NOESY, 3D ^1H - ^{13}C HSQC NOESY y 2D ^1H - ^1H NOESY, se han obtenido las correlaciones NOE, que son los parámetros de RMN que nos proporcionan mayor información estructural. Estas correlaciones nos dan, también, información acerca de la estructura secundaria de la cadena peptídica.

Como se puede observar en la figura 5.11, correspondiente a la intensidad de las señales de correlación NOE del zimógeno ONCFLG, aparecen señales $d_{\alpha N(i,i+1)}$ intensas en las regiones que pertenecen o están en torno a las secuencias de residuos con lámina β y señales $d_{NN(i,i+1)}$ intensas en la regiones de hélice α .

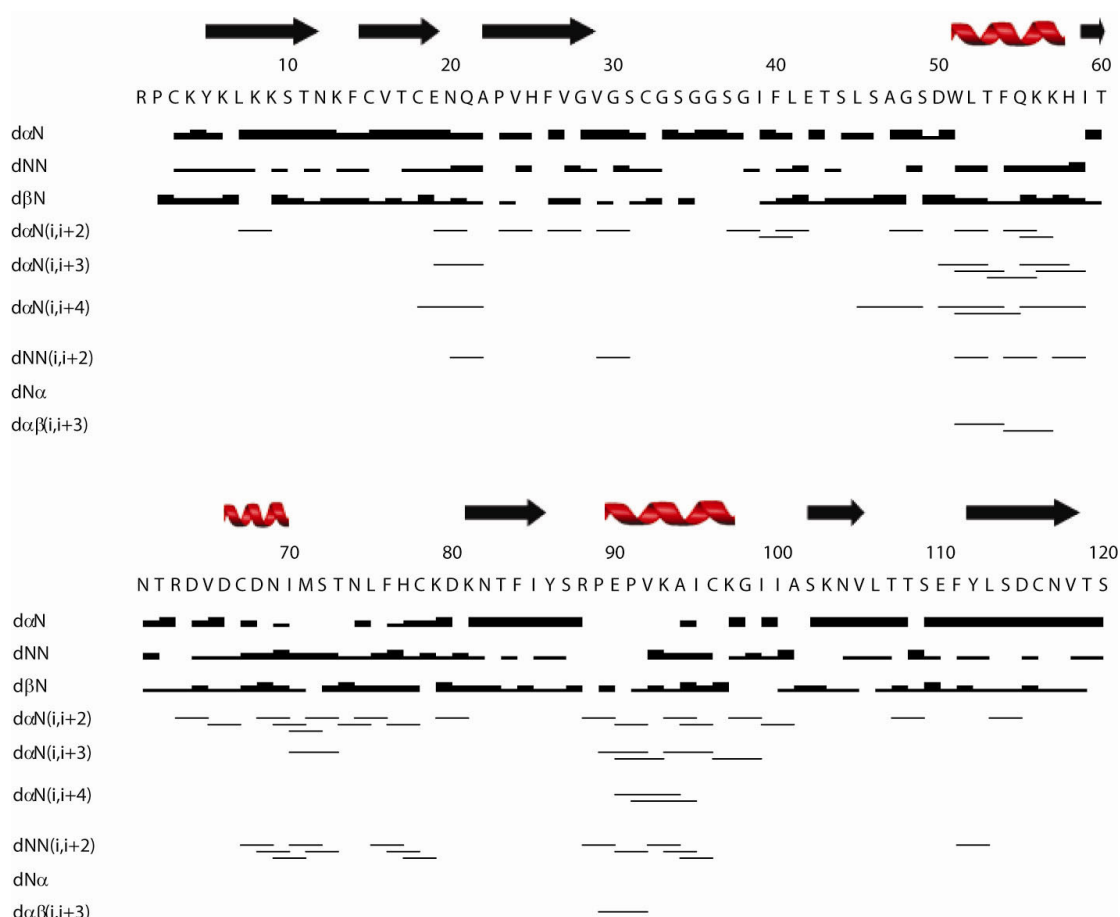


Figura 5.11. Resumen de NOEs secuenciales y de medio alcance del zimógeno ONCFLG. En la parte superior de la secuencia están representados los elementos de estructura secundaria que se han identificado a partir de los desplazamientos conformacionales con flechas negras, las láminas β y espirales rojas, las hélices α .

3.3.2. Estructura tridimensional del zimógeno ONCFLG

La estructura tridimensional se ha calculado siguiendo el procedimiento descrito en el apartado 2.5 de Materiales y Métodos. A partir de las señales obtenidas en los espectros 2D ^1H - ^1H NOESY, 3D ^1H - ^{15}N HSQC NOESY y 3D ^1H - ^{13}C HSQC NOESY se obtuvieron 2025 restricciones de distancia, tabla 5.3. A partir de los desplazamientos químicos de los carbonos $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$ y $^1\text{H}_\alpha$ mediante el programa TALOS+ (Shen et al. 2009), se obtuvieron 149 restricciones de ángulos diedros Φ y ψ y se utilizaron 24 restricciones para definir los enlaces disulfuro.

Restricciones		
Tipo		Nº
De distancia	de largo alcance ($5 \leq i-j $)	1059
	de corto alcance ($ i-j \leq 1$)	288
	de medio alcance ($1 < i-j < 5$)	678
Por enlaces S-S		24
Angulares (Φ, ψ)		149
Número total		2198
Restricciones por residuo		18.3

Tabla 5.3. Número total de restricciones empleadas para el cálculo que incluyen las restricciones de distancia, restricciones angulares y restricciones de los enlaces disulfuro de las Cys.

Como se observa en la figura 5.12 prácticamente todos los residuos, excepto aquellos comprendidos en el puente de unión, entre la Gly33 y Gly48, se caracterizan por tener un número importante de restricciones de largo, medio y corto alcance. La falta de restricciones de distancia para la mencionada región de la proteína es debida a la carencia de estructura ordenada, ya indicada por los desplazamientos conformacionales (figura 5.10; sección 3.3.1).

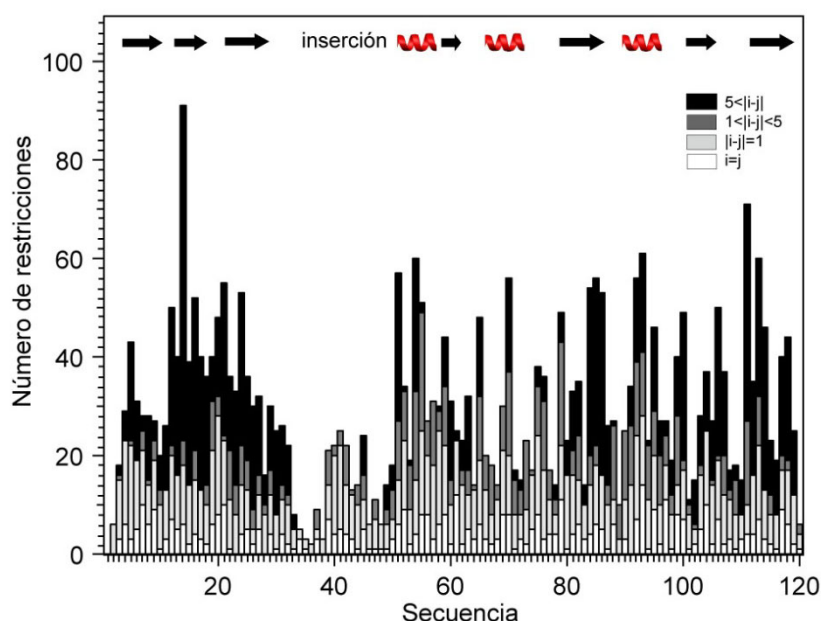


Figura 5.12. Representación del número de restricciones por residuo que se han empleado en el cálculo de la estructura tridimensional del zimógeno ONCFLG. En blanco, se representan las restricciones intrarresiduales; en gris claro, las secuenciales; en gris oscuro, las de medio alcance y en negro, las de largo alcance. En la parte superior están representados los elementos de estructura secundaria, indicando con flechas negras, las láminas β y espirales rojas, las hélices α .

Se calcularon un total de 100 estructuras, considerándose las veinte mejores (menor valor de función blanco) como las que representan la estructura de la proteína en disolución. Los valores obtenidos del cálculo con CYANA se dan en la tabla 5.4.

CYANA (20 estructuras)		
Valor de la función blanco	Mínimo	1.54
	Promedio	1.69
	Máximo	1.96
Violación de las restricciones de distancia máxima (Å)	Mínima	0.12
	Promedio	0.20
	Máxima	0.47
Violación de los ángulos diedros (°)	Mínimo	2.32
	Promedio	3.58
	Máximo	9.51

Tabla 5.4. Valores obtenidos tras el cálculo con CYANA de la función blanco, violaciones de distancia máxima y ángulos diedros.

Todos los ángulos de torsión del esqueleto se encuentran en las regiones permitidas en la representación de Ramachandran, figura 5.13.

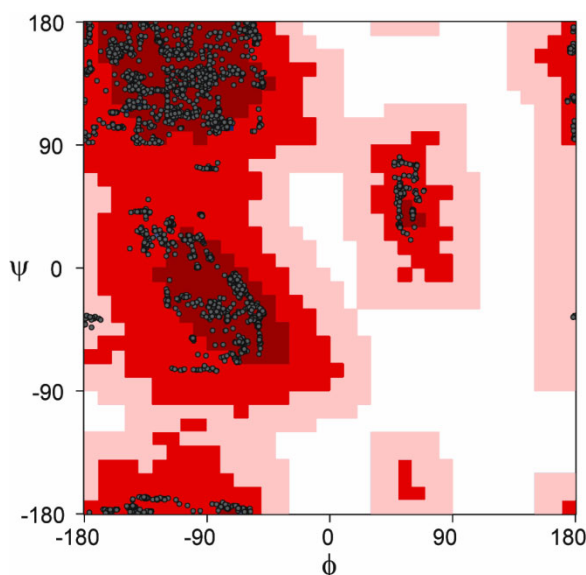


Figura 5.13. Mapa de Ramachandran de los ángulos ϕ y ψ del zimógeno ONCFLG obtenidos con CYANA. El 65.8% de los ángulos se encuentran en regiones favorables, el 33.3% en regiones adicionales permitidas, 0.9% en regiones generosamente permitidas y 0.0% en regiones no permitidas.

La figura 5.14 representa la superposición de las veinte mejores estructuras y se aprecia el plegamiento global del zimógeno ONCFLG. Se puede observar como las regiones del esqueleto con estructura secundaria se superponen perfectamente entre los distintos conformeros, mientras que la superposición de la secuencia insertada presenta más variaciones de un conformero a otro. El zimógeno ONCFLG prácticamente mantiene la misma estructura terciaria que muestran otros miembros de la superfamilia de la RNasa A y es particularmente similar a la Onconasa® silvestre (Mosimann et al. 1994), figura 5.1, sus variantes preparadas por mutagénesis dirigida (Lee et al. 2008; Schulenburg et al. 2010) y la ribonucleasa citotóxica de la rana toro (Chang et al. 1998).

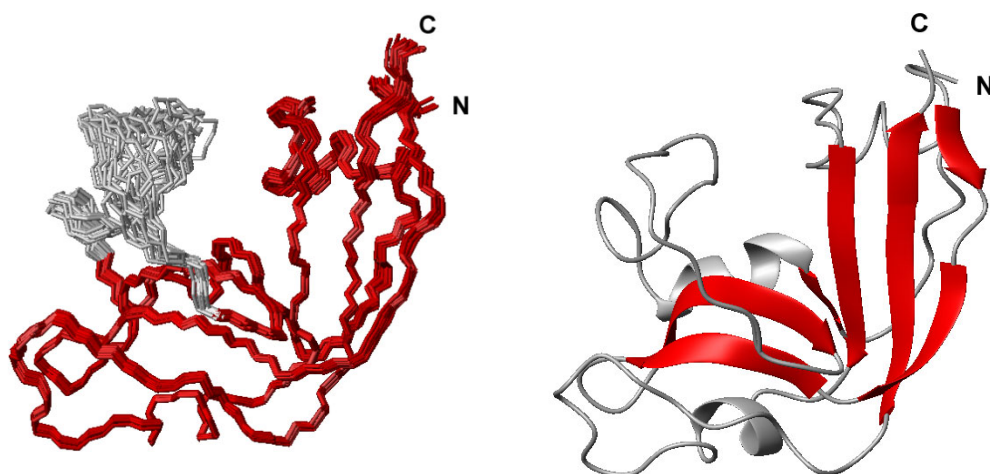


Figura 5.14. Representación de las estructuras finales del zimógeno ONCFLG calculadas con CYANA. A la izquierda, superposición de los átomos del esqueleto de las 20 estructuras finales, la parte en gris representa la inserción de 16 residuos que bloquea el centro activo. A la derecha, representación en cinta de una de las estructuras con las regiones helicoidales en gris y las láminas en rojo.

Una vez analizados los resultados de las estructuras obtenidas con CYANA y considerando que los valores obtenidos están dentro de los límites establecidos se realiza la minimización de energía con AMBER. Primero se realiza un cálculo de control y a continuación el cálculo en vacío y en agua. Los valores obtenidos con AMBER se pueden ver en la tabla 5.5.

AMBER(20 estructuras)	Vacío	En agua
Energía media (Kcal/mol)	-1987.73 (-1611.96 a - 2549.71)	-3926.14 (-3883.71 a -3950.66)
Energía electrostática (Kcal/mol)	-8381.0 (-7983.6 a -9017.2)	-7768.9 (-7038.8 a -8084.0)
Energía de Van der Waals (Kcal/mol)	-756.4 (-731.7 a -771.1)	-822.9 (-813.4 a -834.8)
RMSD enlace de geometría ideal (Å)	0.01 ± 0.001	0.01 ± 0.001
RMSD ángulos de geometría ideal (°)	2.368 ± 0.01	2.171 ± 0.01

Tabla 5.5. Valores obtenidos una vez realizada la minimización energética de las estructuras obtenidas utilizando el programa AMBER.

Las estructuras obtenidas después de la minimización, utilizando el modelo de Born de solvatación, con AMBER9 (Case 2006) y que contiene las coordenadas de los 20 conformeros más estables, son las que se toman como la familia representativa del zimógeno ONCFLG, figura 5.15.

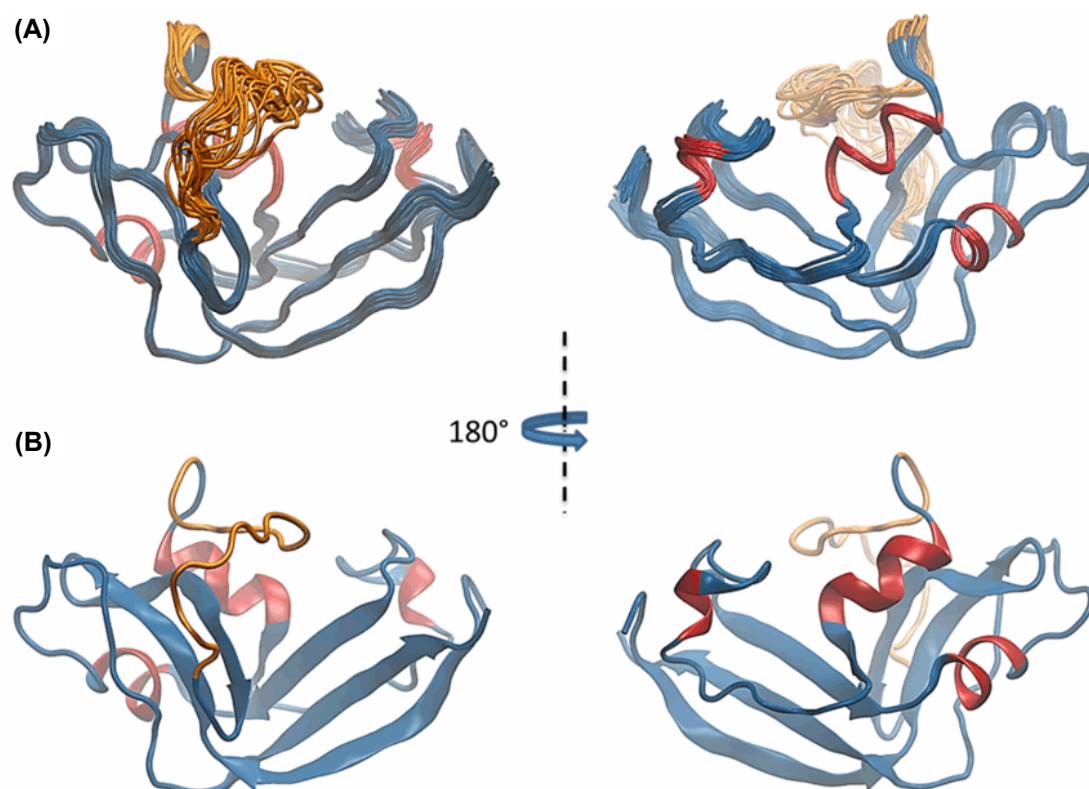


Figura 5.15. (A) Superposición del esqueleto de los 20 conformeros, energéticamente más estables, del zimógeno ONCFLG obtenidos después de minimizar con AMBER. (B) Representación en cinta de uno de los conformeros del zimógeno ONCFLG. El segmento peptídico de unión (Gly33-Gly48) está representado en naranja y el resto del esqueleto de la proteína en azul, excepto las hélices α que están representadas en rojo. Las representaciones están giradas 180° una respecto a la otra. Esta figura ha sido publicada en el artículo (Callís et al. 2013).

Antes de realizar el depósito de la estructura obtenida para el zimógeno en el Protein Data Bank, se ha comprobado la calidad estereoquímica de las estructuras obtenidas del cálculo con CYANA y después de la minimización energética con AMBER mediante la herramienta PROCHECK. En la tabla 5.6 se comparan los valores de los ángulos Φ y ψ en la representación de Ramachandran de las estructuras antes y después de la minimización de energía. Puede verse que el número de residuos en la zona más favorable aumenta tras la minimización de AMBER.

Mapa de Ramachandran (%)	CYANA	AMBER
Residuos en regiones más favorables	66.5	74.9
Residuos en regiones permitidas adicionales	35.4	24.7
Residuos en regiones en general permitidas	1.1	0.0
Residuos en regiones no permitidas	0.0	0.0

Tabla 5.6. Valores en % de los residuos que se encuentran en las distintas regiones del mapa de Ramachandran obtenidos con PROCHECK para los resultados de CYANA y AMBER.

La estructura del zimógeno ONCFLG está formada por tres hélices α y dos láminas β con cuatro y dos hebras respectivamente. Para facilitar la comparación, la estructura secundaria se numera de acuerdo a la topología de la Onconasa®, pdb: ONC1, figura 5.16. Las tres hélices α : α 1, Leu52-His58; α 2, Ile70-Ser72 y α 3, Pro89-Ala94 (numeración del zimógeno) coinciden razonablemente bien con las hélices presentes en la Onconasa® nativa que incluyen los residuos Trp3-His10, Cys19-Ile22 y Pro41-Lys49 en la numeración de la nativa. Las hebras que forman las láminas β , se encuentran en la regiones β 5, Tyr5-Asn12; β 6, Phe14-Cys18; β 7, Pro23-Val29; β 1, Ile59-Thr60; β 2, Lys81-Tyr86; y β 4, Phe111-Val118; y son también prácticamente equivalentes a aquellas de la Onconasa® nativa que se localizan en Tyr77-Asn84, Phe86-Cys90, Pro95-Val101, Ile11-Thr12, Lys33-Tyr38, Phe63-Val70, respectivamente. La diferencia más notable la observamos en el segmento Lys103-Leu106 del zimógeno, que no presenta estructura secundaria definida. Esta zona corresponde con Lys55-Leu58 de la ONC nativa que se ordena dando lugar a la hebra β , β 3, figura 5.16.

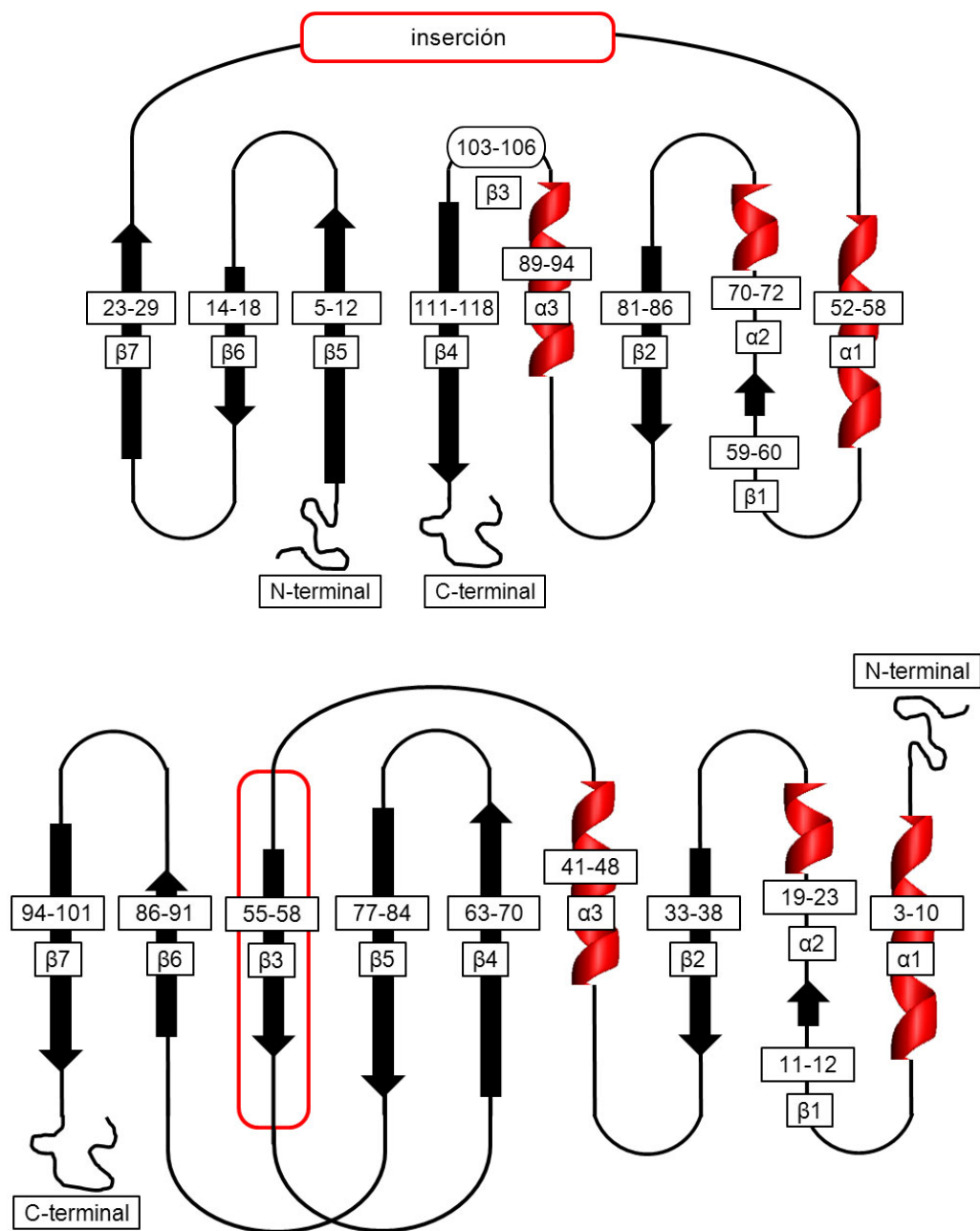


Figura 5.16. Representación esquemática de la estructura secundaria. En la parte superior el zimógeno ONCFLG y en la parte inferior la ONC. La estructura de ambas proteínas es prácticamente idéntica. La diferencia más notable aparece en la Lys55-Leu58 de la ONC, que corresponde a Lys103-Leu106 en ONCFLG. Esta región que no tiene estructura secundaria en el zimógeno forma una hebra β en la Onconasa® nativa.

El RMSD o desviación cuadrática media nos indica la convergencia de los datos experimentales, tabla 5.7. Las estructuras calculadas presentan regiones muy bien definidas que corresponden con aquellas que tienen estructura secundaria, y otras que no lo están tanto y corresponden a regiones sin estructura periódica. El valor de RMSD global para el esqueleto de ONCFLG es de 0.84 Å. Si nos fijamos en el RMSD de las regiones con estructura secundaria, 0.32 Å, podemos afirmar que hay buena convergencia entre las estructuras calculadas, pero ocurre lo contrario si nos fijamos en la región desestructurada, 33-48, con un valor de RMSD de 1.47 Å. Por tanto, la carencia de estructura de la región insertada, hace que el valor global de RMSD de nuestras estructuras sea más elevado. Esta región sobresale del núcleo de la proteína ocupando la cara cóncava de la forma de riñón característica de las RNAsas. De esta manera, el acceso al sitio activo que contiene los residuos catalíticos His25, His58 y Lys79, puede ser impedido por este conector flexible, que puede actuar como un agente bloqueante estérico (Tompa 2002).

RMSD (Å)	
Átomos del esqueleto	
Global (1-120)	0.84±0.31
Región estructurada (1-32,49-120)	0.32±0.07
Región desestructurada (33-48)	1.47±0.61
Estructura secundaria (hélices y láminas)	0.25±0.06
Hélices α (51-58, 70-72, 90-94)	0.21±0.06
Láminas β (5-12, 14-18, 23-29, 59-60, 81-86, 111-118)	0.23±0.05
Átomos pesados	
Global (1-120)	1.31±0.26
Región estructurada (1-32,49-120)	0.96±0.07
Región desestructurada (33-48)	2.13±0.75
Estructura secundaria (hélices y láminas)	0.71±0.09
Hélices α (51-58, 70-72, 90-94)	0.93±0.17
Láminas β (5-12, 14-18, 23-29, 59-60, 81-86, 111-118)	0.55±0.09

Tabla 5.7. Valores de RMSD en angstrom de la familia de estructuras de ONCFLG.

La superficie accesible de la proteína está representada en la figura 5.17. En general, encontramos regiones con muy baja accesibilidad que coinciden mayormente con zonas de estructura secundaria de hebra β , como las comprendidas en los intervalos 14-18, 22-24, 58-59, 82-85 y 113-116. Algunos residuos como Phe54, Ile70, Val92 y Cys96 están prácticamente enterrados y no tienen contacto con el solvente. La zona insertada (residuos 33-48) presenta una exposición elevada de acuerdo con la diversidad estructural, figuras 5.14 y 5.15, y la dinámica que presenta y que se describe más adelante, apartado 3.7.

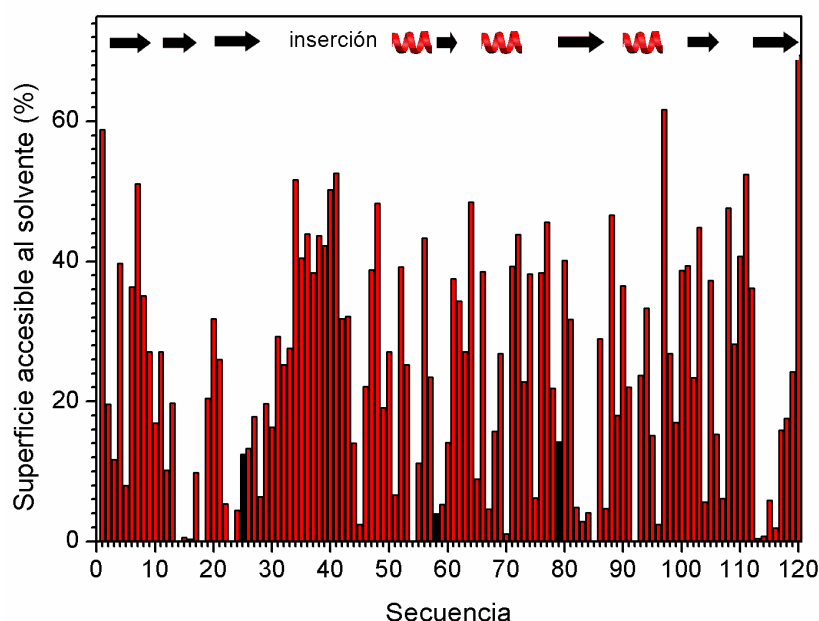


Figura 5.17. Superficie accesible al solvente de cada uno de los residuos que forman la proteína. En la parte superior están representados los elementos de estructura secundaria, indicando con flechas negras, las láminas β y espirales rojas, las hélices α . Se han representado en negro los valores para los residuos del centro activo His25, His58 y Lys79.

Ya se ha comentado que la estructura del zimógeno determinada en este trabajo es similar, en cuanto a plegamiento y estructura secundaria, a la de la proteína de la que procede su diseño. Sin embargo un análisis detallado permite, en este contexto, encontrar algunas diferencias significativas que pudieran tener importancia. Así, en la figura 5.18, podemos observar como ONCFLG presenta una reorganización local sutil en el entorno del enlace disulfuro Cys18-Cys96 con respecto a ONC (Cys48-Cys90). En el zimógeno ONCFLG, hay un incremento de aproximadamente 1 Å en la distancia C_{α} - C_{α} de estas cisteínas. Este incremento está probablemente relacionado con que, en el caso de ONCFLG, la hélice $\alpha 3$ es ligeramente más corta que en ONC.

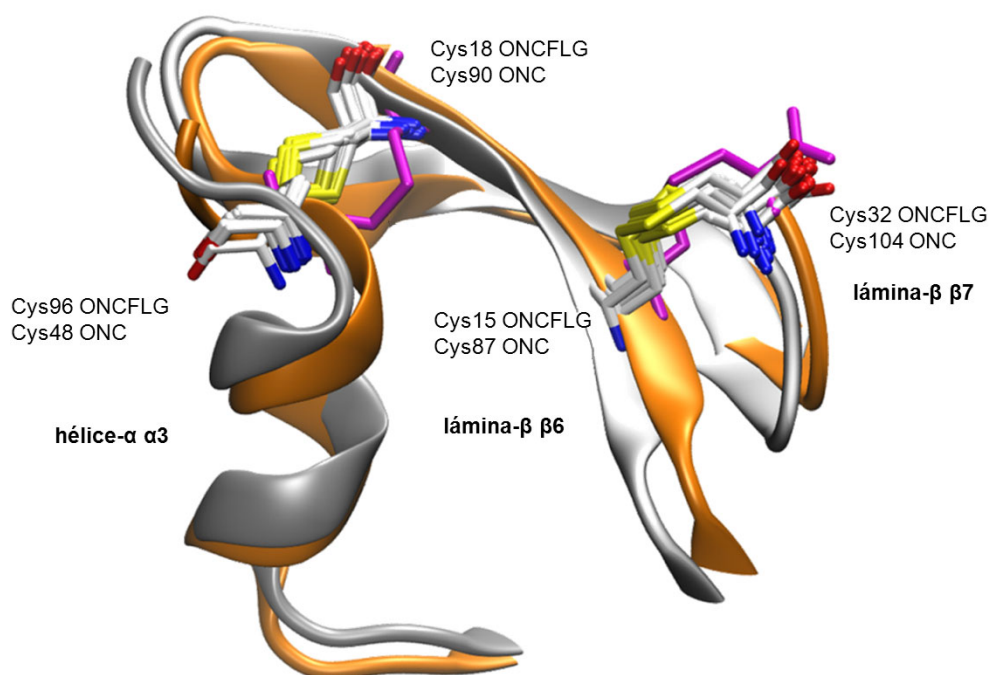


Figura 5.18. Comparación de los enlaces disulfuro análogos de ONCFLG, cuyo esqueleto está representado en gris, y ONC nativa, con el esqueleto en naranja. Los enlaces disulfuro de ONCFLG se muestran en colores CPK, y los de ONC están representados en magenta. Cys18-Cys96 y Cys48-Cys90 unen la lámina β β 6 y la hélice α α 3; y Cys15-Cys32 y Cys87-Cys104 conectan la lámina β β 6 y la lámina β β 7 en ONCFLG y en la ONC (pdb 1ONC), respectivamente. Puede observarse la diferente conformación adoptada por el enlace disulfuro Cys18-Cys96, de ONCFLG, con respecto al Cys48-Cys90, de ONC, que está más a la izquierda en la figura. Esta representación se ha generado usando VMD (Humphrey et al. 1996).

3.3.3. Análisis de la estructura en relación a la actividad ribonucleolítica

Un estudio en profundidad de la estructura del zimógeno ONCFLG, comparada con las estructuras de ONC disponibles, proporciona información acerca de los residuos del sitio activo, otras zonas de unión al RNA y su relación con la baja actividad ribonucleolítica presentada por esta enzima (Callís et al. 2013). En las ribonucleasas, es bien conocida la importancia de los subsitos de unión a RNA en relación a su función (Pares et al. 1991). Estos subsitos se han descrito en la Onconasa® y se conocen como B, cuando implican residuos que interaccionan y reconocen las bases del nucleótido, y como P cuando la interacción es con el grupo fosfato del RNA (Lee et al. 2008). Concretamente, B1 está formado por Lys33, Thr35, Asp67 y Phe98; y B2 por Thr89 y Glu91 (Lee et al. 2008). En la estructura de ONCFLG, todos los residuos pertenecientes a los subsitos B1 (Lys81, Thr83, Asp115 y Phe26) y B2 (Thr17 y Glu19) muestran la misma orientación que los residuos equivalentes en la estructura cristalina de la proteína ONC nativa (código pdb 1ONC)

(Mosimann et al. 1994) o en el complejo de ONC-d(AUGA) (2I5S) (Lee et al. 2008). Sin embargo, se observan diferencias notables para los residuos del sitio activo (o también llamado subsitio P1), His25, His58 y Lys79 y su entorno, figura 5.19.

El O γ de Ser49, que reemplaza al O ϵ del residuo Pyr1 del N-terminal de la proteína nativa, no está formando enlace de hidrógeno con N ϵ de Lys57 (que corresponde a la Lys9 en la ONC nativa) y, en consecuencia, se observa un cambio conformacional de la cadena lateral de Lys57. En la Onconasa®, a Lys31 y a Lys9 se les ha propuesto el papel de estabilizar la acumulación de carga negativa que existe en un oxígeno no unido y fosforilado durante la catálisis (Mosimann et al. 1994; Lee & Raines 2003; Lee et al. 2008). Sin embargo, la cadena lateral de Lys79 en ONCFLG sobresale de la estructura, en vez de apuntar hacia la hendidura del sitio activo como hace su equivalente, Lys31, en la proteína ONC nativa como se muestra en la figura 5.19.

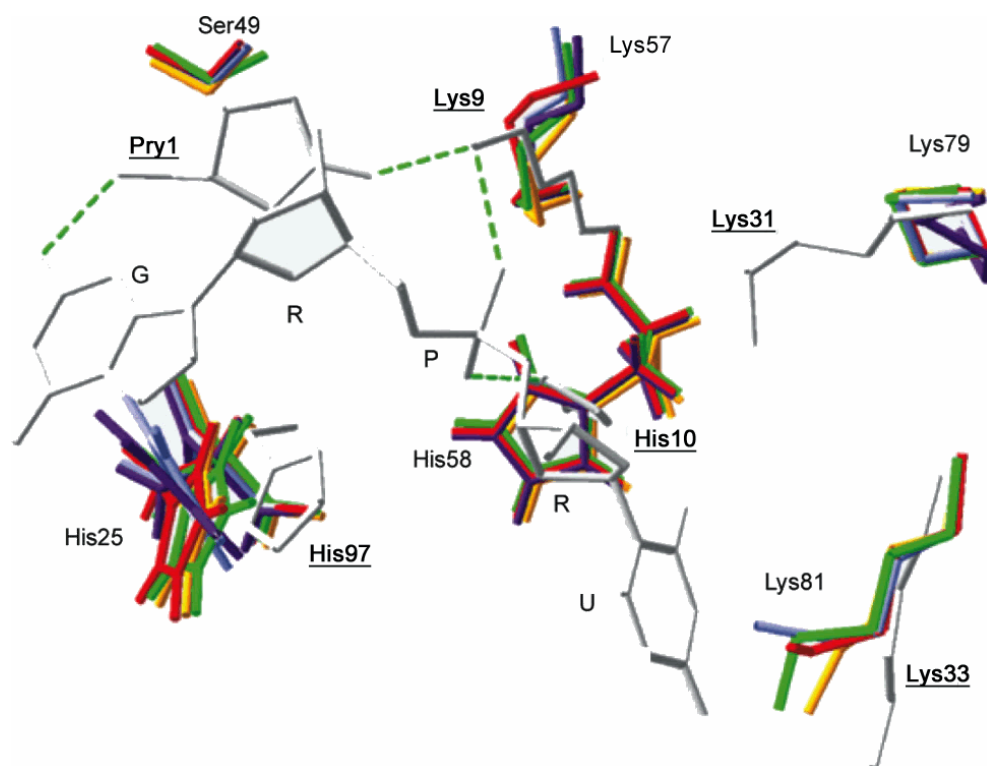


Figura 5.19. Representación de las distintas conformaciones de los residuos del subsitio P1. Los residuos de ONCFLG se muestran en distintos colores y los residuos homólogos de ONC-d(AUGA) están representados en blanco y subrayados. Las letras G, R, P y U indican guanina, ribosa, fosfato y uracilo, respectivamente. Esta imagen ha sido diseñada con el programa Swiss-PdbViewer (Guex and Peitsch 1997). Esta imagen ha sido previamente publicada (Callis et al. 2013).

Como consecuencia del movimiento de Lys57 en la estructura ONCFLG, His58 parece oscilar arriba y abajo de la posición de His10 en la Onconasa®, mostrando una cierta variabilidad conformacional, figura 5.20. En cuanto a His25, la cadena lateral adopta un conjunto de conformaciones en las que los $N_{\delta 1}$ y $N_{\delta 2}$ del anillo imidazólico intercambian su posición a través de la rotación en torno al ángulo diedro $\chi 2$. Además, His25 en ONCFLG está a medio camino entre las dos conformaciones conocidas de la His97. Una de ellas denominada posición A, es la que está presente en el complejo ONC-d(AUGA) (código pdb 2I5S), en amarillo en la figura 5.20 (A). La otra conformación de la His97, llamada posición B, es la que adopta en la estructura de la ONC nativa y que se ha representado en verde en la figura 5.20 (A).

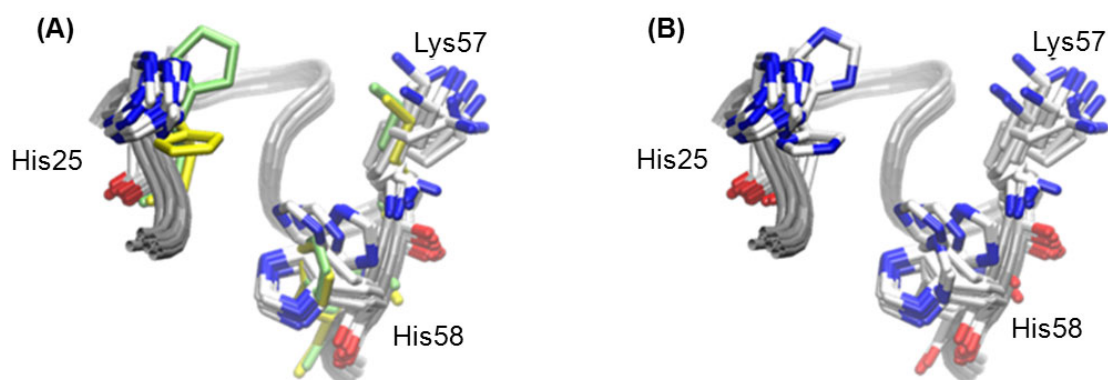


Figura 5.20. Conformaciones adoptadas por His25 y His 58 en la estructura de ONCFLG. (A) Las distintas conformaciones de ONCFLG se muestran en colores CPK, las cadenas laterales de la ONC nativa (1ONC pdb) (His10, Lys31 y His97) en verde y las del complejo ONC-d(AUGA) (2I5S pdb) en amarillo. (B) Representa lo mismo que la figura A excepto que todos los residuos están pintados en CPK. Ambas figuras fueron generadas usando el programa VMD (Humphrey, Dalke et al. 1996). La imagen ha sido publicada en el artículo (Callís et al. 2013).

Uno de los puntos importantes del diseño del zimógeno, y que debe condicionar la baja actividad ribonucleolítica del mismo, es la baja accesibilidad del centro activo al sustrato antes de la activación o hidrólisis. Es interesante por tanto ver cuál es el % de la superficie global accesible de los residuos catalíticos, tabla 5.8.

ONC1	ONCFLG	ONCFLG (hidrolizado)
29.5% His97	12.4% His25	18.8% His25
6.8% His10	3.9% His58	4.8% His58
40.5% Lys31	14.2% Lys79	15.5% Lys79

Tabla 5.8. Superficie accesible de los residuos catalíticos de la Onconasa® nativa, el zimógeno ONCFLG y el zimógeno ONCFLG hidrolizado.

Podemos ver que la accesibilidad de los residuos del centro activo de ONCFLG disminuye bastante con respecto a ONC, a pesar de que los cambios conformacionales observados entre ambas estructuras podrían hacer pensar en un sitio más accesible para ONCFLG. Sin embargo, en este caso y teniendo en cuenta los valores de la tabla 5.8, la presencia de la secuencia insertada realiza adecuadamente su papel y protege el centro activo en ausencia del activador. El problema es que la accesibilidad no mejora sustancialmente después de la hidrólisis, lo que queda patente en la estructura de ONCFLG hidrolizada como se discutirá en el apartado 3.5.

3.4. Unión de ONCFLG a un análogo de sustrato

Para poder evaluar experimentalmente si la disminución de la accesibilidad y las diferencias observadas para los residuos del centro activo de la ONCFLG, condicionan la interacción con el sustrato, se decidió estudiar por RMN la interacción con un análogo no hidrolizable, d[UGG]₃.

Al preparar el posible complejo, inmediatamente después de mezclar d[UGG]₃ con ONCFLG marcado se observó la formación de un precipitado blanco.

Dado que a pH 5.2 ONCFLG está cargado positivamente, y se ven favorecidas las interacciones inespecíficas con d[UGG]₃, se subió el pH para debilitarlas y así favorecer el reconocimiento específico y la unión del oligonucleótido al centro activo. El precipitado persistió hasta pH 6.35, pero se disolvió a pH 7.25, dando una solución transparente en la que poder caracterizar la interacción en el caso de que ésta se haya producido.

Teniendo en cuenta la información que nos proporcionan los experimentos anteriormente realizados y la asignación a pH 5.2 (Serrano et al. 2012), se asignó fácilmente el espectro 2D ¹H-¹⁵N HSQC y la mayoría de picos del 2D ¹H-¹H NOESY del posible complejo. Los espectros 2D ¹H-¹⁵N HSQC a pH 7.25 con y sin d[UGG]₃ son prácticamente idénticos, lo que en principio no apunta a la formación de un complejo específico.

Además, se observaron resultados muy similares para las histidinas de ONCFLG en presencia y ausencia de d[UGG]₃, a pesar de la suposición de que la unión de d[UGG]₃ debería modificar los desplazamientos químicos de los átomos involucrados, especialmente en los residuos del centro activo. En ambos casos, los picos de cruce entre los átomos N_{δ1}, H_{ε1}, N_{ε2} y H_{δ2} de las histidinas en un espectro HSQC en la región de los anillos de imidazol, mostró que la forma predominante en que se encuentra la His77 era la cargada mientras que His58 estaba predominantemente en el estado neutro. Además se determinó que el

tautómero predominante era el H-N_{δ1}, *a priori* el menos favorecido en ausencia de estabilizaciones estructurales (Blomberg et al. 1977). El resultado para la His25 fue menos concluyente, pero sugiere que el anillo de imidazol es neutro, principalmente, y adopta una mezcla de los dos estados tautoméricos posibles bajo estas condiciones. Solo los núcleos de H_α pertenecientes a Phe54 y Gln55, cercanos al centro activo, se han visto ligeramente afectados en presencia del nucleótido. Este resultado contrasta con el observado para la unión de análogos de sustrato a RNasa A. En este caso se perturban fuertemente los desplazamientos químicos de los grupos presentes en el centro activo, incluyendo las His catalíticas (Toiron et al. 1996). Por último, no fue posible identificar de manera inambigua NOEs intermoleculares entre d[UGG]₃ y ¹⁵N-ONCFLG, mientras que entre la RNasa A y sus análogos de sustrato se detectaron del orden de 20 NOEs intermoleculares (Toiron et al. 1996).

Puesto que la unión del sustrato por la Onconasa® se debilita en presencia de sal (Turcotte & Raines 2008), se bajó la fuerza iónica del medio y se repitieron el experimentos de RMN a pH 7.9 y 9.0. En estas condiciones, tampoco se hallaron NOEs intermoleculares inambiguos que indicaran la formación del complejo y sólo se encontraron algunas diferencias significativas en los desplazamientos químicos. Bajo estas condiciones, el anillo de imidazol de His58 continúa en el tautómero H-N_{δ1}. Tanto la His25 como la His77 son neutras a este pH y parecen equilibrarse las poblaciones de los tautómeros H-N_{δ1} y H-N_{ε2}.

3.5. Hidrólisis del zimógeno ONCFLG

Los ensayos realizados por el laboratorio del Dr. Ribó han demostrado que la afinidad de la proteasa por el zimógeno es buena y que la hidrólisis es cuantificable a los 15 minutos de añadir la proteasa y es completa a los 45 minutos de incubación. En la figura 5.21 pueden observarse dos bandas, que representan los dos fragmentos en que se divide la proteína una vez que la proteasa ha escindido el correspondiente enlace. Además, el corte se completa a tiempos de incubación más cortos que en otros zimógenos análogos y utilizando relaciones molares proteasa/zimógeno pequeñas, del orden de 1/100. Este resultado pone de manifiesto que la cadena polipeptídica de unión basada en la secuencia de aminoácidos GSGIF*LETSL es cortada de manera eficiente por la proteasa VIH-1.

Siguiendo este protocolo se ha hidrolizado la muestra de ONCFLG para su estudio estructural por RMN.

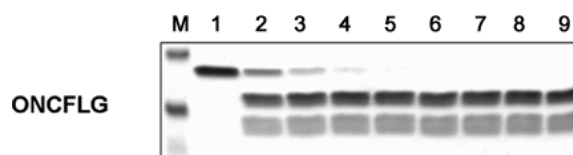
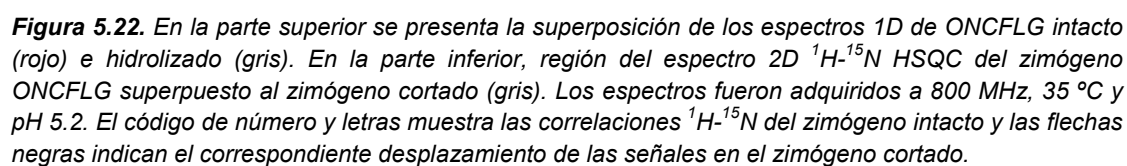


Figura 5.21. Hidrólisis del zimógeno ONCFLG por la proteasa VIH-1. La hidrólisis fue medida a distintos tiempos una vez que se había añadido la proteasa en una proporción proteasa/zimógeno 1/5. La línea M se observan las franjas correspondientes a los patrones de 15, 10 y 4 kDa. Las líneas de la 1-9 corresponden a las muestras tomadas a 0 min, 15 min, 30 min, 45 min, 1 h, 1.30 h, 2h, 20 h y 30 h. Descrito en el apartado de Materiales y Métodos (Callís et al. 2013).

3.6. Estructura del zimógeno ONCFLG hidrolizado

Para determinar los cambios estructurales producidos por la acción de proteasa de VIH-1 sobre el zimógeno ONCFLG, se llevó a cabo la asignación de las resonancias del zimógeno hidrolizado siguiendo la misma metodología descrita para el zimógeno intacto. La figura 5.22 muestra una comparación de los espectros de RMN de ambas construcciones. Los cambios después de la hidrólisis son sutiles y apenas se aprecian en los espectros 1D RMN. Sin embargo, tal y como cabría esperar, en el ^1H - ^{15}N HSQC se observan algunos cambios claros, que vamos a describir a continuación. Dado que la proteasa del VIH-1 debe hidrolizar en el sitio de reconocimiento el enlace Phe40-Leu41, la señal correspondiente al H_N de la Leu41 no aparece en el zimógeno hidrolizado ya que se ha convertido en un grupo amina que intercambia rápido con el solvente. La mayoría de las señales H_N aparecen a valores de desplazamientos químicos similares a los del zimógeno no hidrolizado, aunque un número de residuos tienen cambios significativos. Entre ellos destaca Phe40 adyacente al sitio de corte de la proteasa VIH-1.



Las diferencias entre los desplazamientos químicos de los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ ($\Delta\delta$) entre el zimógeno ONCFLG hidrolizado y sin hidrolizar se encuentran representadas en la figura 5.23. La mayoría de las diferencias son pequeñas, por debajo de 0.1 ppm, excepto las del segmento Phe40-Asp50 localizado cerca del enlace hidrolizado. Es curioso que estas diferencias sean mayores en la nueva parte N-terminal generada a partir de la hidrólisis en 40-41, que en la anterior a esta posición. Este hecho ha sido también observado en péptidos cortos, lo que podría indicar que la carga positiva del extremo N-terminal afecta en mayor medida los desplazamientos químicos que la del carboxilato C-terminal (Jimenez 1987). Los datos de los desplazamientos químicos muestran claramente la conservación de la estructura global de ONCFLG después de la hidrólisis y que las diferencias se localizan localmente cerca de la escisión del enlace Phe40-Leu41.

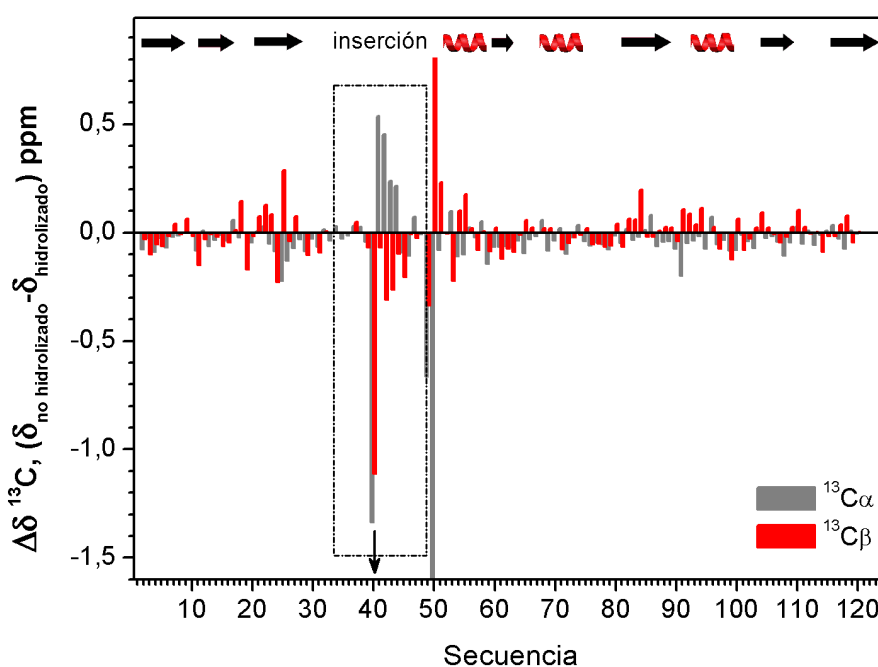


Figura 5.23. Comparación de los valores de los desplazamientos químicos de $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ para las variantes, hidrolizada y no hidrolizada, del zimógeno ONCFLG. Los valores $\Delta\delta$ son la diferencia de desplazamientos químicos entre el zimógeno no hidrolizado e hidrolizado ($\Delta\delta = \delta_{\text{no hidro.}} - \delta_{\text{hidro.}}$). La flecha marca la posición del enlace péptido hidrolizado y las líneas discontinuas el segmento peptídico de unión, Gly33-Gly48. La estructura secundaria está representada en la parte superior de la figura, láminas- β como flechas negras y hélices- α como espirales en rojo.

Por otra parte, el análisis del espectro NOESY nos permitió identificar un conjunto similar de NOEs, de corto, medio y largo alcance, en las regiones estructuradas del zimógeno hidrolizado y sin hidrolizar. Esto apunta a que sus estructuras tridimensionales deben ser muy similares. Utilizando esta información, se construyó un modelo estructural para el zimógeno

ONCFLG hidrolizado que se encuentra representado en la figura 5.24. En este modelo queda de manifiesto el elevado número de conformaciones disponibles para los segmentos que bordean el sitio de escisión mientras que se mantiene la estructura plegada del resto de la secuencia. Estas secuencias flexibles, de alguna manera siguen protegiendo al centro activo puesto que la superficie accesible de los residuos catalíticos no recuperan los valores de la ONC como sería deseable e indican los valores de la tabla 5.8. Así, todos los datos de RMN disponibles para esta construcción, muestran la elevada flexibilidad de estos segmentos, incluyendo la dinámica que se describe a continuación.

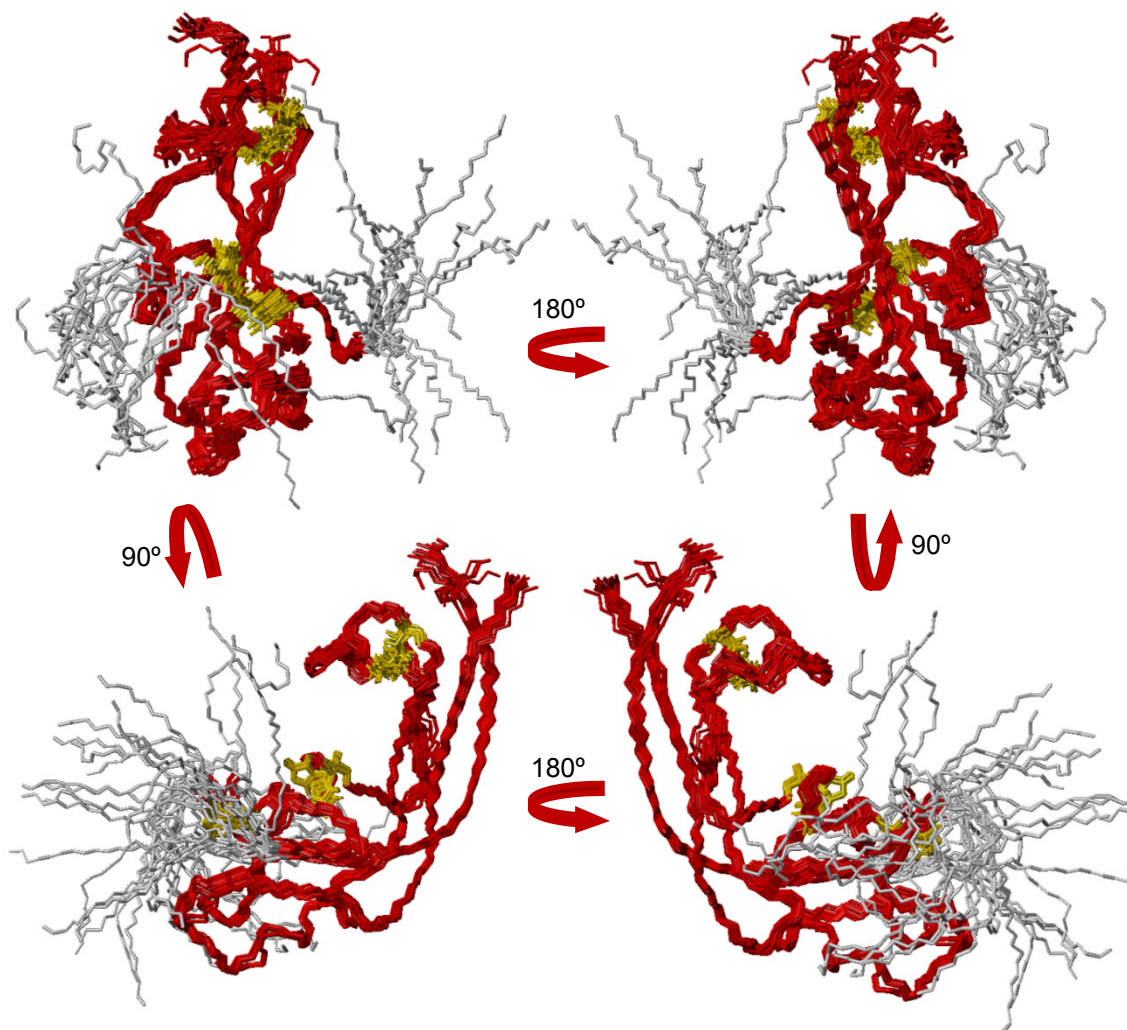


Figura 5.24. Modelo de la estructura del zimógeno ONCFLG hidrolizado. Los residuos del sitio activo, de His25, His58 y Lys79 están representados en amarillo, la secuencia de unión (Gly33-Gly48) en gris mientras el resto de la proteína en rojo. Esta figura se ha generado con el programa MOLMOL (Koradi et al. 1996).

3.7. Dinámica del zimógeno hidrolizado y no hidrolizado

Para poder determinar la flexibilidad local del esqueleto del zimógeno ONCFLG en la escala de tiempo de ns-ps, se han medido los NOEs heteronucleares, figura 5.25. La mayoría de residuos muestran unos valores de NOE cercanos a los valores teóricos esperados para vectores ^{15}N - ^1H rígidos, de acuerdo con los resultados reportados previamente para la Onconasa® y sus variantes (Chang et al. 1998; Gorbatyuk et al. 2004).

Los valores de NOE heteronuclear indican la restricción del enlace N-H al movimiento, por tanto, valores más bajos indican menor restricción, es decir, mayor libertad de movimientos de los enlaces involucrados. Tanto en el zimógeno hidrolizado como sin hidrolizar encontramos que los extremos N- y C- terminales muestran valores bajos de NOE (0.5-0.6) lo que nos indica la flexibilidad de los mismos.

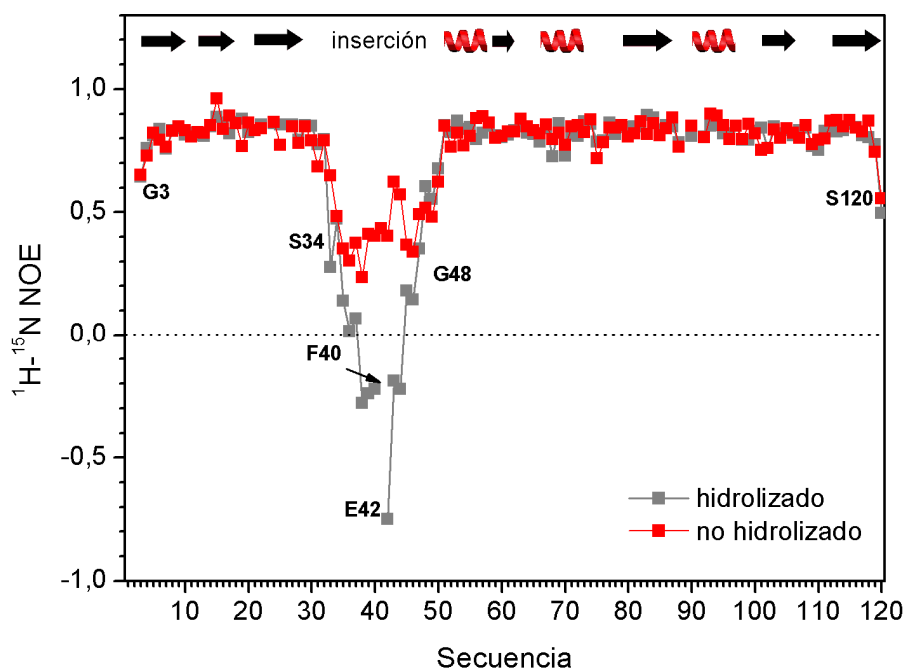


Figura 5.25. Dinámica de las variantes, hidrolizada y no hidrolizada, del zimógeno ONCFLG. Valor de NOE ^1H - ^{15}N para el zimógeno ONCFLG no hidrolizado (rojo) e hidrolizado (gris) en función al número de secuencia. Las líneas usadas para conectar los cuadrados rojos o grises sirven sólo de guía. La estructura secundaria está representada en la parte superior de la figura, láminas- β como flechas negras y hélices- α como espirales rojas. La flecha indica el lugar de la hidrólisis.

Los valores de NOE para la región insertada, de Gly33 a Gly48, en el caso del zimógeno sin hidrolizar muestran un valor medio de NOE de 0.38, lo que sugiere que la región se comporta como un bucle flexible con movimientos rápidos en la escala de tiempo de los ns-ps, característicos de las regiones desestructuradas. En el zimógeno hidrolizado, llegan a

alcanzarse valores de NOE negativos lo que indica que la libertad conformacional de los nuevos extremos, es muy próxima a la de péptidos cortos desestructurados (Kay 1998) con valores de NOE de signo distinto al de las proteínas globulares. Por tanto, los valores de NOE heteronuclear indican, como cabía esperar, que la flexibilidad de la secuencia insertada se incrementa enormemente en la forma hidrolizada del zimógeno mientras que permanece prácticamente idéntica en el resto de la proteína.

3.8. Estabilidad conformacional

La alta estabilidad conformacional supone una ventaja para la proteínas, en concreto para las proteínas con posibles fines terapéuticos ya que una mayor estabilidad supone menor susceptibilidad a la degradación proteolítica no específica, o a cambios con la temperatura. El punto medio del desplegamiento térmico, determinado por DSC del zimógeno ONCFLG es de 80.5 °C (Callís et al. 2013), considerablemente superior a la temperatura fisiológica que es de aproximadamente 37 °C y comparable a los valores determinados para la Onconasa® nativa, 87.5 °C (Callís et al. 2013) y que coinciden con los reportados en la literatura (Notomista et al. 2001). Esto nos permite asegurar que estos zimógenos son lo suficientemente estables para ser utilizados tanto *in vitro* como *in vivo* bajo condiciones fisiológicas.

Se midió la estabilidad conformacional, ΔG_{HX} , para cada uno de los residuos del zimógeno ONCFLG a través de intercambio de H/D monitorizado por RMN. Como observamos en la figura 5.26, los grupos amida más protegidos y estables, están localizados en las regiones de estructura secundaria, principalmente en las láminas β y la primera y tercera hélices α .

Tal como ocurre con la mayoría de proteínas, incluida la RNasa A (Neira et al. 1999), el promedio de ΔG_{HX} para los elementos de la estructura secundaria de la ONCFLG es más alto que en el resto de la proteína. Hay que mencionar que el segmento insertado, no muestra protección frente al intercambio H/D bajo estas condiciones, lo que está de acuerdo con sus propiedades dinámicas. Los protones amídicos de la segunda hélice también están desprotegidos, tal y como ha sido descrito para la ONC y algunas variantes (Gorbatyuk et al. 2004). En la RNasa A y la proteína humana RNasa 3 (Laurents et al. 2009) la hélice que ocupa la posición correspondiente también muestra una estabilidad baja, lo que sugiere que la geometría de los enlaces de hidrógeno formados por esta hélice no es la óptima.

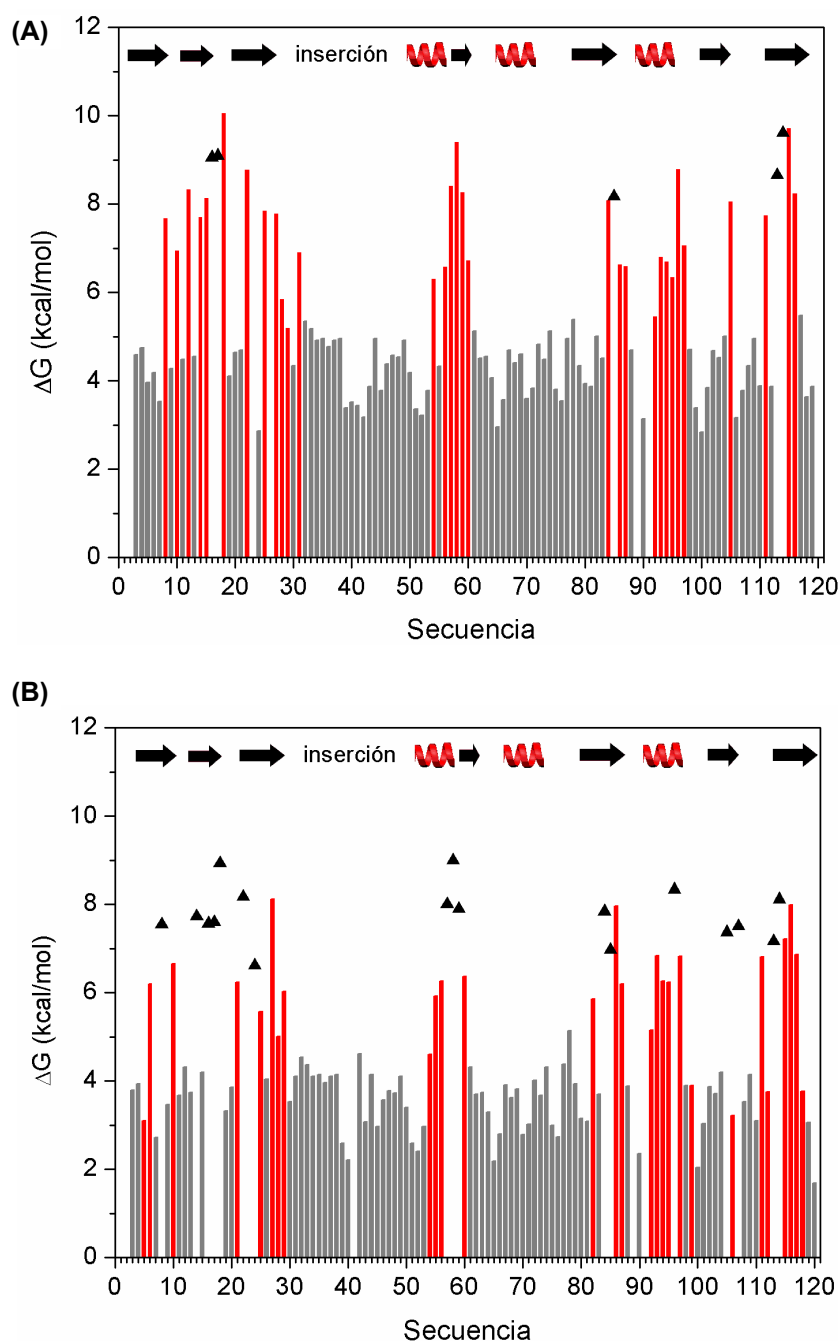


Figura 5.26. Valores ΔG_{HX} obtenidos por RMN del zimógeno ONCFLG sin hidrolizar (A) e hidrolizado (B) en función a la secuencia. Las barras grises representan los límites máximos para el intercambio rápido que no puede ser observado. Las barras rojas muestran los valores medidos experimentalmente y los triángulos indican el límite más bajo estimado para los grupos con intercambio lento. La estructura secundaria está representada en la parte superior de la figura, láminas β como flechas negras y hélices- α como espirales rojas.

Tras la hidrólisis, los valores de ΔG_{HX} son menores en promedio. Estos valores se han utilizado para el cálculo de la T_d del zimógeno hidrolizado que no se ha podido medir experimentalmente de manera sencilla. Asumiendo que la dependencia de la energía libre de estabilidad en el valor de T_d observada para la ONC silvestre (Notomista et al. 2000) se mantiene para el zimógeno ONCFLG, podemos estimar que la hidrólisis por la proteasa podría disminuir su T_d en alrededor de 6.6 °C hasta un valor de 73.9 °C. Esta T_d está todavía muy por encima de 37 °C, por lo que la ONCFLG hidrolizada se puede mantener perfectamente plegada a temperaturas fisiológicas. Este resultado es importante para una futura aplicación biomédica.

4. DISCUSIÓN

Los fundamentos del diseño y la producción del zimógeno de la ribonucleasa A (Plainkum et al. 2003) especifican que la secuencia peptídica que conecta los extremos N- y C- terminal de la enzima nativa, debe bloquear y extenderse sobre el acceso al centro activo para poder tener la funcionalidad que se pretende. El zimógeno debe ser activado cuando una proteasa reconoce y rompe el motivo específico presente en esta secuencia. Por tanto, el segmento puente ha de ser suficientemente largo para conectar los extremos sin inducir tensión y contener el motivo de reconocimiento de la proteasa, pero lo suficientemente corto para que bloquee el sitio activo y así evitar que el sustrato de RNA se una al zimógeno no activado.

Probablemente, la Onconasa® no se había considerado previamente como un pilar en el que apoyarse para el diseño de zimógenos de ribonucleasa a causa de la bien conocida contribución del residuo N-terminal de piroglutamato a la estabilidad y, más particularmente, a su baja actividad catalítica y elevada citotoxicidad (Mosimann et al. 1994; Notomista et al. 2001; Lee & Raines 2003). Sin embargo, la Onconasa®, es por sí misma citotóxica y se ha alcanzado la fase III de ensayos clínicos para el tratamiento del mesotelioma maligno (Goparaju et al. 2011; Nasu et al. 2011). Por esto, la creación de zimógenos eficaces basados en la Onconasa® pondría a disposición de la comunidad científica una nueva y versátil opción para dirigir la actividad enzimática y orientar su citotoxicidad a una célula en un estado de enfermedad específico. Los zimógenos basados en la Onconasa® activados por proteasas microbianas o virales podrían extender la utilidad terapéutica de la Onconasa® para distintas enfermedades, además del cáncer, y representan una estrategia para evitar los mecanismos conocidos de resistencia patógena.

El pequeño tamaño de la Onconasa® sugiere, *a priori*, que una secuencia puente de pequeño tamaño sería suficiente para rellenar la hendidura del centro activo mientras une los extremos originales. Sin embargo, la disposición espacial de la Cys104 terminal, la cual forma un enlace disulfuro con Cys87, requiere que la secuencia de unión insertada deba rodear parte de la molécula. Mediante modelado molecular, se dedujo que una secuencia peptídica de unión de 14 aminoácidos como mínimo cumplía estos requisitos. Por este motivo fueron diseñados, en el laboratorio del Prof. Ribó, varias generaciones de zimógenos que diferían en la longitud y flexibilidad de esta secuencia insertada con el fin de optimizar el reconocimiento por parte de la proteasa y la catálisis (Callís et al. 2013). Por otra parte, Beck y colaboradores (Beck et al. 2000) habían reportado previamente que el sustrato peptídico GSGIF*LETSL se escindía, en el enlace marcado con el asterisco, 60 veces más eficazmente que el natural basado en la secuencia SGVSQNY*PIVQVL. Así pues, se generaron las variantes de

zimógeno de ONC hasta llegar a la variante optimizada denominada ONCFLG, que es la que se ha estudiado en este trabajo.

El grupo del Prof. Ribó ha demostrado que la hidrólisis de ONCFLG por la proteasa del VIH-1 es muy efectiva. De hecho, la reacción se completa a relaciones molares proteasa-zimógeno pequeñas (del orden de 1/100) y con tiempos de incubación muy cortos (menos de 1h) (Callís et al. 2013). Sin embargo, para alcanzar los objetivos fijados al inicio de este trabajo hace falta cumplir otros requisitos. En este sentido, los resultados obtenidos de la actividad ribonucleolítica de la forma hidrolizada de ONCFLG no son satisfactorios (Callís et al. 2013). Si bien la actividad del zimógeno intacto es muy reducida con respecto a la Onconasa modelo tal como era deseado, una vez hidrolizado la eficiencia catalítica únicamente aumenta 1.16 veces (Callís et al. 2013), lo que está muy lejos de lo que se había esperado.

Para poder avanzar en el diseño futuro de formas mejoradas que presenten una buena activación al ser hidrolizadas, quedan por resolver varias cuestiones la mayoría relacionadas con el papel de la secuencia de unión, tanto en la forma intacta como en la hidrolizada. Un primer paso es entender porque el rendimiento enzimático de la proteína supuestamente activada, es sorprendentemente tan bajo. En este sentido hay que preguntarse: ¿Cómo es la estructura de este diseño que lleva insertada la secuencia de enlace?, ¿cómo interacciona el conector con la estructura principal de la Onconasa®, ¿cómo su conformación puede afectar a la orientación final de los residuos que flanquean el enlace escindible y a los residuos del centro activo?, etc. El análisis de las estructuras tridimensionales de las formas hidrolizada e intacta de ONCFLG que hemos descrito debe arrojar luz a estos interrogantes.

4.1. Estructura de la secuencia insertada en el zimógeno ONCFLG

La estructura tridimensional del zimógeno ONCFLG resuelta por espectroscopía de RMN, está bien definida y es bastante similar a la de la Onconasa® nativa, excepto en la secuencia enlazante, Gly33-Gly48. A pesar de que la modelización previa que se llevó a cabo predecía que esta secuencia podía adoptar una estructura de hélice α , figura 5.5, la estructura resuelta demuestra que, una vez insertada, no adopta ninguna estructura secundaria regular y muestra una gran diversidad estructural, figuras 5.14 y 5.15. Además, los datos de NOE heteronuclear, indican claramente que este segmento de unión es también la región más dinámica de la molécula incluso en la forma no hidrolizada, figura 5.25.

La flexibilidad conformacional exhibida por la secuencia enlazante, junto con la accesibilidad de los residuos Phe40-Leu41 que contiene el sitio de corte, figuras 5.14 y 5.15,

explican la escisión rápida y eficaz del zimógeno ONCFLG por la proteasa específica. En este sentido, el diseño de nuevos zimógenos mejorados debe respetar estas propiedades de accesibilidad que favorecen la activación.

Las medidas de intercambio H/D a nivel de residuo corroboran la extraordinaria estabilidad previamente determinada por DSC. La T_d del zimógeno está solo unos grados por debajo de la que presenta la Onconasa®. Esta notable estabilidad es algo sorprendente considerando que los extremos N- y C- terminales se han permutado en una secuencia natural. La secuencia peptídica insertada no muestra ningún tipo de protección frente al intercambio H/D, tanto en el zimógeno hidrolizado como en el zimógeno sin hidrolizar, lo que es consistente con la falta de una estructura secundaria bien definida y el comportamiento altamente dinámico de esta región indicado por los resultados de relajación, figura 5.26. Tras el corte, los valores de ΔG_{HX} de toda la proteína son ligeramente menores en promedio lo que sugiere que existen interacciones entre el segmento peptídico de unión y el centro de la proteína y que estas contribuyen a la estabilidad conformacional global. Sin embargo estas interacciones no han podido ser confirmadas por los datos de RMN ya que no se ha detectado ningún NOE entre la secuencia insertada y el resto de la proteína más allá de los NOEs secuenciales en los extremos de conexión.

4.2. Bases estructurales de la actividad del zimógeno ONCFLG

La comparación de la estructura del zimógeno ONCFLG con las estructuras disponibles de la Onconasa® nativa también arroja luz sobre la variación de la actividad ribonucleolítica observada (Callís et al. 2013). Aunque no existen diferencias en la posición de los residuos pertenecientes a los subsitios de reconocimiento de las bases, B1 y B2, se observan cambios importantes a nivel de los residuos catalíticos en el subsitio P1, figura 5.19. Raines y colaboradores (Lee & Raines 2003) mostraron que la sustitución de Pyr1 por Ala provocaba una disminución de 29 veces de la actividad ribonucleolítica de la ONC y sugirió que la pérdida del enlace de hidrógeno entre O_ϵ de Pyr1 y N_ζ de Lys9 proporcionaba libertad de rotación a Lys9, reduciendo la catálisis. Además, el cambio de Lys9 o Lys31 por una Ala reducía el valor de K_{cat}/K_M en 103 veces. La actividad catalítica del mutante doble (Lys9Ala, Lys31Ala) estaba por debajo del límite de la sensibilidad del ensayo. Estos resultados llevan a la conclusión de que tanto Lys9 como Lys31 contribuyen a la estabilización del estado de transición.

La estructura de ONCFLG revela la pérdida del enlace de hidrógeno homólogo al mencionado en el párrafo anterior. Esta pérdida puede ser determinante para la libertad conformacional observada en la cadena lateral de Lys57 (Lys9 en la ONC nativa) y apoya la

idea de que la función de Pyr1 en la Onconasa® es posicionar apropiadamente la cadena lateral de Lys9 para la catálisis. En el caso de Lys79 (Lys31 en la ONC), la cadena lateral de la misma se aleja del cuerpo principal de ONCFLG, figura 5.19. Es probable que esta conformación le impida cumplir el papel descrito para Lys31 en ONC.

En ONCFLG, hemos visto que la movilidad de His58 aumenta y el anillo de imidazol de His25 puede girar alrededor del ángulo diedro χ_2 sin adoptar ninguna de las conformaciones definidas como A o B, observadas en el complejo ONC-d(AUGA) (2I5S) y en la ONC nativa (1ONC), figuras 5.19 y 5.20. Estos cambios en las histidinas catalíticas pueden ser consecuencia de los movimientos de Lys57 y de la presencia de la secuencia insertada que une los extremos N- y C- terminales de la secuencia natural, y probablemente afectan a la actividad de la enzima.

Por otra parte y tal como era deseable, hemos comprobado que la presencia del enlazador limita el acercamiento de cualquier sustrato para la catálisis, figura 5.27, y obstaculiza el correcto posicionamiento y unión de d[UGG]₃ en la hendidura del sitio activo. El pequeño número de perturbaciones de los desplazamientos químicos y la ausencia de NOEs intermoleculares entre el oligonucleótido estudiado y ONCFLG son los datos que apoyan este hecho.

En conjunto, todas estas características estructurales y dinámicas, que se derivan de la presencia de la secuencia peptídica de unión intacta, justifican razonablemente bien la disminución de la actividad ribonucleolítica de ONCFLG con respecto a la ONC (eficiencia catalítica 7 veces menor).

Finalmente, nuestro grupo colaborador de Girona ha demostrado que la ONCFLG no hidrolizada se internaliza bien y no muestra citotoxicidad alguna en los linfocitos T y células Jurkat, lo cual es una característica muy deseable para un zimógeno. Además, la internalización del zimógeno ONCFLG en células Jurkat se ha seguido por microscopia confocal y se ha visto que a partir de las 3-6 horas de incubación ya es detectado dentro de la célula (Callís et al. 2013). Todo esto va en la línea correcta de lo que se había mencionado en la introducción de este capítulo, ya que para ser efectivos los zimógenos diseñados deben tener baja actividad enzimática, no producir citotoxicidad y presentar buena internalización antes de la activación.

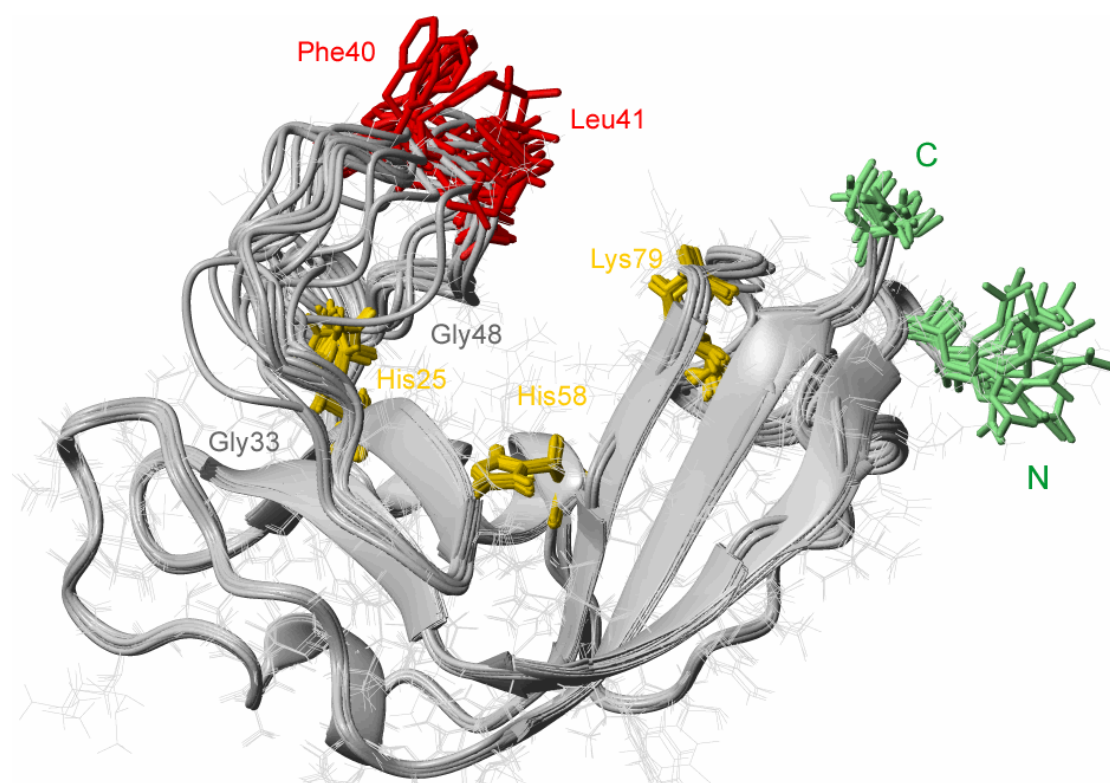


Figura 5.27. Estructura del zimógeno ONCFLG no hidrolizado. En la región Gly33-Gly48, que carece de estructura secundaria definida encontramos los residuos Phe40-Leu41, que constituye el sitio de corte que se muestra en rojo. Los residuos del centro activo, His58, Lys79 y His25, están representados en amarillo y en verde los extremos N- y C- terminales. Ésta representación ha sido generada usando el programa MOLMOL (Koradi et al. 1996).

4.3. Bases estructurales de la actividad del zimógeno ONCFLG hidrolizado

Tal y como se ha comentado a lo largo del capítulo, los datos de RMN (desplazamientos químicos, tautómeros de His y NOEs) apuntan a que la estructura global y específicamente la del centro activo se mantiene intacta en ONCFLG después de la hidrólisis. Por ello cabe suponer que las modificaciones en la estructura del centro activo detalladas y discutidas en el apartado anterior se mantienen en el ONCFLG hidrolizado.

Por otra parte, el modelo estructural de esta molécula, figura 5.24, revela que los segmentos en que queda dividida la inserción peptídica pueden adoptar un gran número de conformaciones después de la hidrólisis. El movimiento aleatorio de estos segmentos podría obstruir estéricamente la aproximación del sustrato al sitio activo, aumentar la K_M y empeorar la ya pobre afinidad de unión de nucleótidos que justifica la baja eficiencia ribonucleolítica de ONC (Leland et al. 1998; Lee & Raines 2003). Este efecto, junto con las alteraciones en los

residuos del sitio activo del zimógeno descritas anteriormente, podría ser la causa de la baja actividad catalítica del zimógeno modificado con respecto a la ONC. Esto es un problema a mejorar en posteriores diseños ya que lo óptimo hubiera sido observar un aumento sustancial de la eficiencia catalítica después del tratamiento con la proteasa y llegar a recuperar la que presenta la Onconasa® nativa.

Los datos estructurales que se han descrito en este capítulo, sientan las bases para una futura optimización del diseño de zimógenos de Onconasa® a través del rediseño de la longitud y la secuencia de la cadena enlazante que permita mejorar la posición de los grupos aceptores de los enlaces de hidrógeno para fijar N_ζ de Lys57, o encaminadas a insertar sustituciones adicionales como Met71Leu (Met28Leu en la ONC (Merlino et al. 2005) destinadas a aumentar la actividad catalítica, o el diseño de sitios adicionales de escisión de la proteasa dentro de la secuencia de unión para disminuir el impedimento estérico de los segmentos flexibles después de la escisión.

5. CONCLUSIONES

En este trabajo se ha caracterizado la estructura tridimensional, la dinámica y la estabilidad de un zimógeno de la Onconasa® que es hidrolizado por la proteasa del VIH-1, y se han relacionado estos datos con la actividad catalítica y las posibles mejoras del diseño para las aplicaciones farmacológicas de estas construcciones en terapias dirigidas.

La estructura 3D del zimógeno ONCFLG determinada por RMN está bien definida, mantiene el plegamiento general de las RNasas y en concreto se parece a la Onconasa® de donde procede su diseño.

Se han detectado modificaciones locales con respecto a la geometría de la ONC nativa. Concretamente el enlace disulfuro Cys18-Cys96 adopta una conformación diferente a su homólogo Cys48-Cys90. Esta diferencia podría estar relacionada con la menor longitud que adopta la hélice α_3 en ONCFLG.

El centro activo de ONCFLG presenta variaciones en la conformación de las cadenas laterales de las histidinas His25 e His48 con respecto a la ONC. El anillo de His25 está en una posición intermedia entre las definidas como A y B en las RNasas, y la His48 presenta cierta variabilidad conformacional. Las cadenas laterales de Lys57 y Lys79 también difieren en su conformación de la adoptada en ONC.

El segmento de unión Gly33-Gly48 no adopta en el zimógeno ninguna estructura secundaria definida, mantiene una elevada flexibilidad lo que parece favorable para la hidrólisis del enlace Phe40-Leu41 por la proteasa específica del VIH-1 y protege al centro activo disminuyendo su accesibilidad.

La estabilidad conformacional medida por RMN está de acuerdo con las determinaciones calorimétricas que indican que ONCFLG es una molécula muy estable, $T_m > 70$ °C, que se mantiene plegada por encima de la temperatura fisiológica, lo que es beneficioso para un futuro terapéutico de los zimógenos basados en Onconasa®.

La ausencia de variaciones en los desplazamientos químicos y de NOEs intermoleculares indican que la ONCFLG no se une a un análogo de sustrato no hidrolizable.

A partir de los datos de RMN, se ha construido un modelo del zimógeno hidrolizado. Los fragmentos generados por la hidrólisis del segmento de unión, son altamente flexibles y adoptan conformaciones aleatorias que producen un impedimento estérico en la zona del centro activo que dificulta la aproximación del sustrato. Este hecho junto a las modificaciones

conformacionales de los grupos catalíticos pueden ser la razón de la baja actividad del zimógeno activado.

A partir de los datos aquí reportados se proponen tres alternativas para mejorar el diseño de estas moléculas de manera que puedan tener una aplicación terapéutica en el futuro: (I) fijar la cadena lateral de Lys57 para que mejore la interacción con el fosfato del sustrato; (II) insertar sustituciones adicionales como Met71Leu que aumenta la actividad catalítica en la ONC y (III) insertar sitios adicionales de hidrólisis de la proteasa en el conector para disminuir el efecto estérico de la cadena escindida y exponer el centro activo de la molécula activada.

6. BIBLIOGRAFÍA

- Ardelt W, Mikulski SM, Shogen K (1991) Amino acid sequence of an anti-tumor protein from *Rana pipiens* oocytes and early embryos. Homology to pancreatic ribonucleases. The Journal of biological chemistry 266, 245-251.
- Arnold U, Ulbrich-Hofmann R (2006) Natural and engineered ribonucleases as potential cancer therapeutics. Biotechnology letters 28, 1615-1622.
- Bai Y, Milne JS, Mayne L et al. (1993) Primary structure effects on peptide group hydrogen exchange. Proteins 17, 75-86.
- Barnard EA (1969) Biological function of pancreatic ribonuclease. Nature 221, 340-344.
- Barre-Sinoussi F, Chermann JC, Rey F et al. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220, 868-871.
- Beck ZQ, Hervio L, Dawson PE et al. (2000) Identification of efficiently cleaved substrates for HIV-1 protease using a phage display library and use in inhibitor development. Virology 274, 391-401.
- Blomberg F, Maurer W, Ruterjans H (1977) Nuclear magnetic resonance investigation of ¹⁵N-labeled histidine in aqueous solution. Journal of the American Chemical Society 99, 8149-8159.
- Boix E, Wu Y, Vasandani VM et al. (1996) Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity. Journal of molecular biology 257, 992-1007.
- Callís M, Serrano S, Benito A et al. (2013) Towards Tricking a Pathogen's Protease into Fighting Infection: The 3D Structure of a Stable Circularly Permuted Onconase Variant Cleaved by HIV-1 Protease. PloS one 8.
- Case DA, Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S. and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.
- Cavanagh J, Fairbrother WJ, Palmer III AG et al. (1996) Protein NMR spectroscopy: Principles & Practice. Academic Press.
- Chang CF, Chen C, Chen YC et al. (1998) The solution structure of a cytotoxic ribonuclease from the oocytes of *Rana catesbeiana* (bullfrog). Journal of molecular biology 283, 231-244.
- Costanzi J, Sidransky D, Navon A et al. (2005) Ribonucleases as a novel pro-apoptotic anticancer strategy: review of the preclinical and clinical data for ranpirnase. Cancer investigation 23, 643-650.
- Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. Journal of biomolecular NMR 13, 289-302.
- Darzynkiewicz Z, Carter SP, Mikulski SM et al. (1988) Cytostatic and cytotoxic effects of Pannon (P-30 Protein), a novel anticancer agent. Cell and tissue kinetics 21, 169-182.
- Dauer B (2005) Protease inhibitors: the current status. Journal of HIV therapy 10, 72-74.
- Debouck C (1992) The HIV-1 protease as a therapeutic target for AIDS. AIDS research and human retroviruses 8, 153-164.
- Dickson KA, Haigis MC, Raines RT (2005) Ribonuclease inhibitor: structure and function. Progress in nucleic acid research and molecular biology 80, 349-374.
- Duerst R, Verbach A (1965) Accurate NMR temperature Measurements. Review of Scientific Instruments 36, 1896-1897.
- Fyfe CA (1983) Solid State NMR for Chemists. C.F.C. Press.

- Goddard TD, Kneller DG (2005) Sparky 3, University of California, San Francisco.
- Goparaju CM, Blasberg JD, Volinia S et al. (2011) Onconase mediated NFKbeta downregulation in malignant pleural mesothelioma. *Oncogene* 30, 2767-2777.
- Gorbatyuk VY, Tsai CK, Chang CF et al. (2004) Effect of N-terminal and Met23 mutations on the structure and dynamics of onconase. *The Journal of biological chemistry* 279, 5772-5780.
- Güntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353-378.
- Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* 14, 33-38, 27-38.
- Jimenez MA (1987) Estructura secundaria en disolución de fragmentos peptídicos de Ribonucleasa A. Tesis doctoral.
- Johnson RJ, Lin SR, Raines RT (2006) A ribonuclease zymogen activated by the NS3 protease of the hepatitis C virus. *The FEBS journal* 273, 5457-5465.
- Kay LE (1998) Protein dynamics from NMR. *Nature structural biology* 5 Suppl, 513-517.
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *Journal of molecular graphics* 14, 51-55, 29-32.
- Laskowski RA, Rullmannn JA, MacArthur MW et al. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR* 8, 477-486.
- Laurents DV, Bruix M, Jimenez MA et al. (2009) The (1)H, (13)C, (15)N resonance assignment, solution structure, and residue level stability of eosinophil cationic protein/RNase 3 determined by NMR spectroscopy. *Biopolymers* 91, 1018-1028.
- Lee JE, Bae E, Bingman CA et al. (2008) Structural basis for catalysis by onconase. *Journal of molecular biology* 375, 165-177.
- Lee JE, Raines RT (2003) Contribution of active-site residues to the function of onconase, a ribonuclease with antitumoral activity. *Biochemistry* 42, 11443-11450.
- Leland PA, Schultz LW, Kim BM et al. (1998) Ribonuclease A variants with potent cytotoxic activity. *Proceedings of the National Academy of Sciences of the United States of America* 95, 10407-10412.
- Liao YD, Wang SC, Leu YJ et al. (2003) The structural integrity exerted by N-terminal pyroglutamate is crucial for the cytotoxicity of frog ribonuclease from *Rana pipiens*. *Nucleic acids research* 31, 5247-5255.
- Merlino A, Mazzarella L, Carannante A et al. (2005) The importance of dynamic effects on the enzyme activity: X-ray structure and molecular dynamics of onconase mutants. *The Journal of biological chemistry* 280, 17953-17960.
- Mikulski SM, Costanzi JJ, Vogelzang NJ et al. (2002) Phase II trial of a single weekly intravenous dose of ranpirnase in patients with unresectable malignant mesothelioma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 20, 274-281.
- Mosimann SC, Ardelt W, James MN (1994) Refined 1.7 Å X-ray crystallographic structure of P-30 protein, an amphibian ribonuclease with anti-tumor activity. *Journal of molecular biology* 236, 1141-1153.
- Nasu M, Carbone M, Gaudino G et al. (2011) Ranpirnase Interferes with NF-kappaB Pathway and MMP9 Activity, Inhibiting Malignant Mesothelioma Cell Invasiveness and Xenograft Growth. *Genes & cancer* 2, 576-584.
- Neira JL, Sevilla P, Menendez M et al. (1999) Hydrogen exchange in ribonuclease A and ribonuclease S: evidence for residual structure in the unfolded state under native conditions. *Journal of molecular biology* 285, 627-643.
- Notomista E, Catanzano F, Graziano G et al. (2000) Onconase: an unusually stable protein. *Biochemistry* 39, 8711-8718.
- Notomista E, Catanzano F, Graziano G et al. (2001) Contribution of chain termini to the conformational stability and biological activity of onconase. *Biochemistry* 40, 9097-9103.

- Pares X, Nogues MV, de Llorens R et al. (1991) Structure and function of ribonuclease A binding subsites. *Essays in biochemistry* 26, 89-103.
- Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of biomolecular NMR* 2, 661-665.
- Plainkum P, Fuchs SM, Wiyakrutta S et al. (2003) Creation of a zymogen. *Nature structural biology* 10, 115-119.
- Ribo M, Bosch M, Torrent G et al. (2004) Quantitative analysis, using MALDI-TOF mass spectrometry, of the N-terminal hydrolysis and cyclization reactions of the activation process of onconase. *European journal of biochemistry / FEBS* 271, 1163-1171.
- Rodriguez M, Torrent G, Bosch M et al. (2007) Intracellular pathway of Onconase that enables its delivery to the cytosol. *Journal of cell science* 120, 1405-1411.
- Rutkoski TJ, Raines RT (2008) Evasion of ribonuclease inhibitor as a determinant of ribonuclease cytotoxicity. *Current pharmaceutical biotechnology* 9, 185-189.
- Saxena SK, Gravell M, Wu YN et al. (1996) Inhibition of HIV-1 production and selective degradation of viral RNA by an amphibian ribonuclease. *The Journal of biological chemistry* 271, 20783-20788.
- Saxena SK, Sirdeshmukh R, Ardelt W et al. (2002) Entry into cells and selective degradation of tRNAs by a cytotoxic member of the RNase A family. *The Journal of biological chemistry* 277, 15142-15146.
- Schulenburg C, Low C, Weininger U et al. (2009) The folding pathway of onconase is directed by a conserved intermediate. *Biochemistry* 48, 8449-8457.
- Schulenburg C, Weininger U, Neumann P et al. (2010) Impact of the C-terminal disulfide bond on the folding and stability of onconase. *Chembiochem : a European journal of chemical biology* 11, 978-986.
- Serrano S, Callis M, Vilanova M et al. (2012) (1)H, (13)C and (15)N resonance assignments of the Onconase FL-G zymogen. *Biomolecular NMR assignments*.
- Shen Y, Delaglio F, Cornilescu G et al. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* 44, 213-223.
- Souza DG, Bertini R, Vieira AT et al. (2004) Repertaxin, a novel inhibitor of rat CXCR2 function, inhibits inflammatory responses that follow intestinal ischaemia and reperfusion injury. *British journal of pharmacology* 143, 132-142.
- Toiron C, Gonzalez C, Bruix M et al. (1996) Three-dimensional structure of the complexes of ribonuclease A with 2',5'-CpA and 3',5'-d(CpA) in aqueous solution, as obtained by NMR and restrained molecular dynamics. *Protein science : a publication of the Protein Society* 5, 1633-1647.
- Tomasselli AG, Heinrikson RL (2000) Targeting the HIV-protease in AIDS therapy: a current clinical perspective. *Biochimica et biophysica acta* 1477, 189-214.
- Tomba P (2002) Intrinsically unstructured proteins. *Trends in biochemical sciences* 27, 527-533.
- Turcotte RF, Raines RT (2008) Design and characterization of an HIV-specific ribonuclease zymogen. *AIDS research and human retroviruses* 24, 1357-1363.
- Whitehead B, Craven CJ, Waltho JP (1997) Double and triple resonance NMR methods for protein assignment. *Methods Mol Biol* 60, 29-52.
- Wishart DS, Bigam CG, Holm A et al. (1995a) (1)H, (13)C and (15)N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of biomolecular NMR* 5, 332.
- Wishart DS, Bigam CG, Holm A et al. (1995b) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of biomolecular NMR* 5, 67-81.
- Wu Y, Mikulski SM, Ardelt W et al. (1993) A cytotoxic ribonuclease. Study of the mechanism of onconase cytotoxicity. *The Journal of biological chemistry* 268, 10686-10693.
- Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. J Wiley&Sons, Inc New York.

Wüthrich K (1990) Protein structure determination in solution by NMR spectroscopy. The Journal of biological chemistry 265, 22059-22062.

Capítulo VI

Resumen y Conclusiones Generales

Capítulo VI:

Resumen y Conclusiones Generales

El trabajo expuesto en esta memoria se ha llevado a cabo en el Grupo de Estructura, Dinámica e Interacciones de Proteínas por RMN del Departamento de Química Física Biológica del Instituto de Química Física Rocasolano del Consejo Superior de Investigaciones Científicas durante el período de Noviembre de 2009 hasta Noviembre de 2013.

En cada uno de los capítulos de esta memoria, que recogen los resultados experimentales, se han incluido las conclusiones concretas de los sistemas abordados y las técnicas utilizadas. A modo de resumen, incluimos aquí una serie de conclusiones generales que pueden deducirse de este trabajo centrado en las interacciones proteína-membrana, que se han estudiado por RMN a partir de modelos simplificados de diversos sistemas concretos.

1. La RMN en disolución es una técnica muy versátil para el estudio de sistemas modelo relacionados con interacciones proteína-membrana. Desde el punto de vista metodológico, todavía hay lugar para el desarrollo de nuevas secuencias de pulsos, como se hace en el capítulo II. Esta nueva secuencia diseñada, puede ser ventajosa no sólo para el estudio de proteínas intrínsecamente desordenadas antes y después de las interacciones proteína-membrana, sino también para el estudio de proteínas globulares que pierden la estructura como consecuencia de la interacción.

2. Desde el punto de vista aplicado, la complejidad de los sistemas proteína-membrana hace necesaria la utilización de modelos simplificados: péptidos para simular regiones de proteínas, y micelas y solventes orgánicos para simular entornos de membrana. De forma general, esta aproximación funciona bien y la sensibilidad es buena gracias a los avances que ha experimentado la RMN en disolución en los últimos años.

3. Los resultados obtenidos en modelos simplificados de interacciones proteína-membrana pueden extrapolarse al sistema biológico concreto, lo que permite postular mecanismos o explicar comportamientos en el entorno fisiológico (caps. IV y V).

4. Dada la complejidad de las interacciones proteína-membrana y que el conocimiento actual sobre ellas es aún escaso, las hipótesis de partida no siempre se cumplen. En este sentido, hemos comprobado que, a pesar de la similitud previamente postulada entre los derivados de heparina y los lipopolisacáridos (LPS), éstos interactúan de manera completamente diferente con los péptidos boomerang (cap. III). Por otra parte, el diseño del zimógeno OCFLG parecía óptimo para los fines terapéuticos que se deseaban, pero ni la estructura ni la recuperación de la actividad catalítica cumplieron las expectativas (cap. V).

5. El trabajo en medios micelares es complejo y difícil, y la estabilidad de las muestras y la calidad de los espectros de RMN obtenidos no siempre es la idónea para llevar a cabo un trabajo estructural. De hecho, hay dos temas que han tenido que abandonarse o posponerse por estos motivos a lo largo de esta Tesis. El primero es el estudio de los péptidos boomerang en presencia de LPSs. El segundo es la caracterización de la Onconasa® en medios micelares de SDS y DPC. A pesar de que la asignación espectral en ambos casos se ha abordado, por el momento no ha sido posible obtener información detallada de lo que sucede en estas interacciones.

6. El hecho de que todos los sistemas estudiados presenten variabilidad estructural en los entornos de membrana indica que la flexibilidad conformacional, y no sólo la estructura, es un requisito indispensable para la funcionalidad biológica de los complejos sistemas proteína-membrana.

Los resultados obtenidos se recogen en 4 artículos ya publicados y otros dos que se encuentran en fase de redacción:

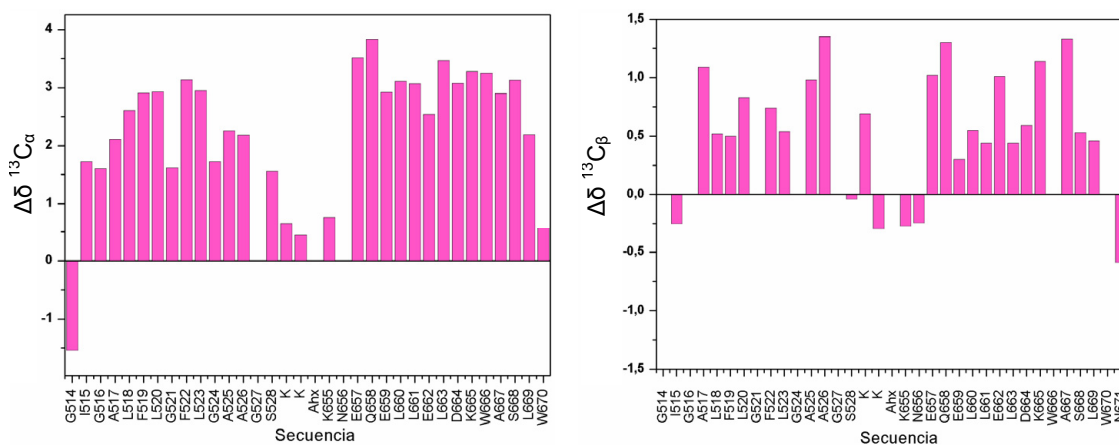
- Callís M., Serrano S., Benito A., Laurents D.V., Vilanova M., Bruix M. and Ribó M. *"Towards Tricking a Pathogen's Protease into Fighting Infection: The 3D Structure of a Stable Circularly Permuted Onconase Variant Cleaved by HIV-1 Protease"* PLoS One 8 **(2013)** (8) e54568.
- Serrano S., Callís M., Vilanova M., Benito A., Laurents D.V., Ribó M., Bruix M. *"(1)H, (13)C and (15)N resonance assignments of the Onconase FL-G zymogen"* Biomol. NMR Assign. **(2013)** 7,13-15.
- Bermel W., Bruix M., Felli I.C., Kumar V., Pierattelli R. and Serrano S. *"Improving the chemical shift dispersion of multidimensional NMR spectra of intrinsically disordered proteins"* J. Biomol. NMR **(2013)** 55, 231-237.
- Serrano S., Araujo A., Apellániz B., Bryson S., Carravilla P., De la Arada I., Huarte N., Rujas E., Pai E.F., Arrondo J.L., Domene C., Jiménez M.A., Nieva J.L. *"Structure and immunogenicity of a peptide vaccine including the complete HIV-1 gp41 2F5 epitope. Implications for antibody recognition mechanism and immunogen design"* J. Biol. Chem. **(2014)**.

Apéndice

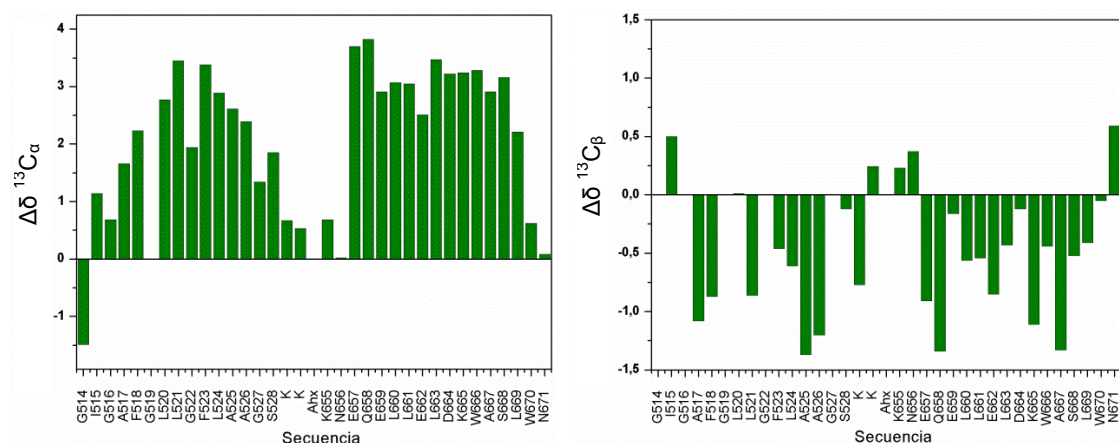
Apéndice

FIGURA A.1.

Diagramas de barras de los valores de $\Delta\delta C_\alpha$ y $\Delta\delta C_\beta$ del péptido HybK3 en 25% HFIP. Los valores positivos para los C_α y negativos para los C_β son representativos de hélice α .



Diagramas de barras de $\Delta\delta C_\alpha$ y $\Delta\delta C_\beta$ del péptido scrHybK3 en 25% HFIP.



Nota: El ensanchamiento que se produce en los espectros en DPC impide que hayamos podido determinar los valores de los desplazamientos químicos de los carbonos de ninguno de los dos péptidos en este medio.

FIGURA A.2.

Diagramas de barras de los $\Delta\delta C_\alpha$ y $\Delta\delta C_\beta$ del péptido CpreTM en 25% HFIP.

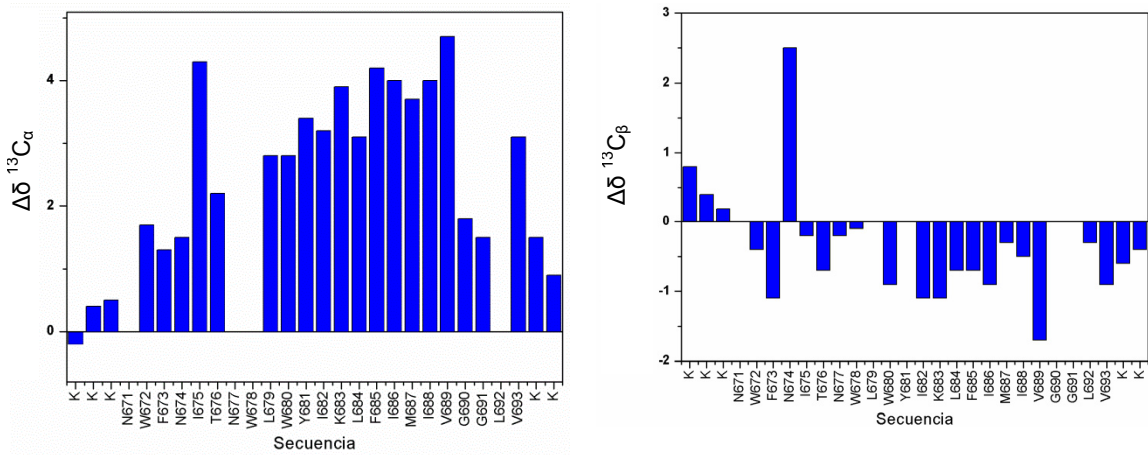


FIGURA A.3.

Diagrama de barras de los desplazamientos químicos conformacionales de $^1H_\alpha$ ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{random coil}}$, ppm) en función de la secuencia para los péptidos HybK3, MPERp, CpreTM y en 20 mM DPC. Los valores que han sido empleados como referencia sin estructura son los de Wishart et al (Wishart et al. 1995a; Wishart et al. 1995b).

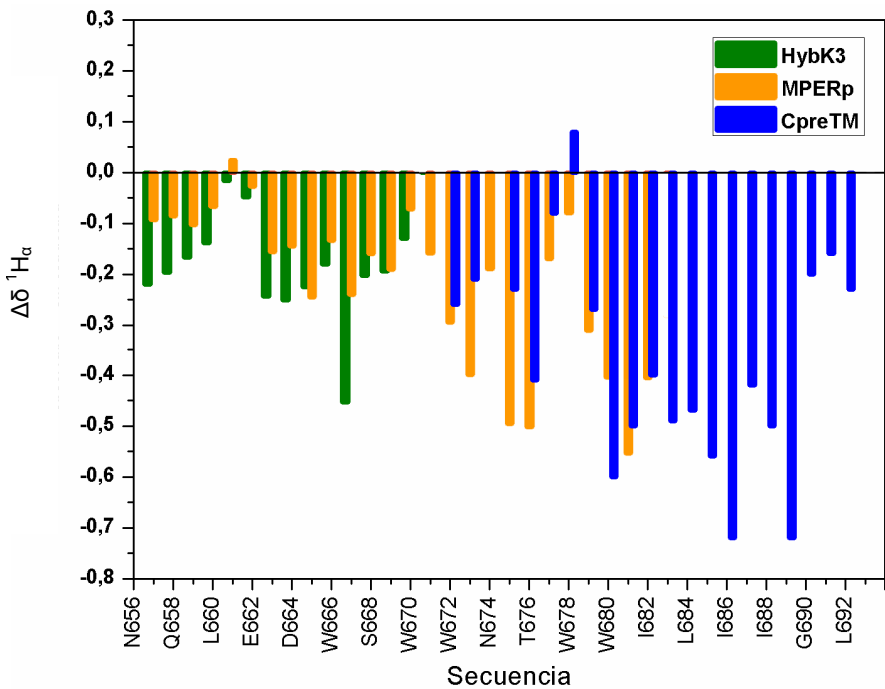


TABLA A.1. Desplazamientos químicos de ^1H , ^{13}C y ^{15}N para el péptido **HybK3** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **25% HFIP** a 25 °C.

Residuo	HN	CaH	CβH	Otros
Gly 514		3.86, 3.94 (43.56)		
Ile 515	8.22	4.18 (62.83)	1.93 (39.05)	C _{V1} H 1.28, 1.51 (27.80) C _{V2} H ₃ 0.98 (17.02) C _δ H ₃ 0.92 (12.69)
Gly 516	8.28	3.83, 3.91 (46.71)		NH 109.48
Ala 517	7.70	4.18 (54.61)	1.47 (18.00)	NH 121.32
Leu 518	7.57	4.17 (57.71)	1.65, 1.77 (41.88)	C _V H (27.44) C _{δ1} H ₃ 0.88 (22.93) C _{δ2} H ₃ 0.95 (24.39) NH 118.11
Phe 519	7.76	4.43 (60.61)	3.20 (39.10)	C _δ H 7.17 C _ε H 7.26 NH 118.54
Leu 520	7.98	4.10 (58.03)	1.56 1.84 (41.57)	C _V H 1.80 (27.14) C _δ H ₃ 0.92 (22.35) NH 118.62
Gly 521	7.91	3.81 (46.72)		NH 106.36
Phe 522	7.97	4.36 (60.84)	3.21 (38.86)	C _δ H 7.14 NH 122.69
Leu 523	8.23	3.94 (58.05)	1.49, 1.77 (41.86)	C _V H 1.63 (26.73) C _{δ1} H ₃ 0.84 (22.42) C _{δ2} H ₃ 0.84 (24.35) NH 120.32
Gly 524	8.20	3.82 (46.83)		NH 105.72
Ala 525	7.88	4.18 (54.76)	1.47 (18.12)	NH 124.76
Ala 526	8.31	4.07 (54.68)	1.29 (17.75)	NH 121.53
Gly 527	8.25	3.87, 3.95		NH 109.65
Ser 528	7.87	4.38 (59.86)	4.02 (63.84)	NH 115.07
Lys 529	7.83	4.34 (56.86)	1.93 (32.41)	C _δ H 1.73 C _V H 1.48, 1.52 NH 121.33
Lys 530	7.92	4.29 (56.65)	1.81, 1.86 (33.40)	C _δ H 1.72 (28.92) C _V H 1.48 C _ε H 3.021 (42.38) NH 120.58
Ahx	7.52	2.33, 2.37 (38.10)	1.66 (27.30)	C _δ H 1.55 (30.83) C _ε H 3.21 (42.19) C _V H 1.35 (28.53) NH 120.64
Lys 655	8.061	4.35 (56.96)	1.81, 1.89 (33.37)	C _δ H 1.76 (29.13) C _ε H 3.04 (42.20) C _V H 1.50 (24.75) NH 124.14
Asn 656	8.11	4.69	3.00 (39.14)	N _{δ2} H 6.74, 7.38 (111.43) NH 115.80

Glu 657	9.09	4.06 (60.12)	2.13 (28.88)	C _v H 2.42, 2.46 (35.57) NH 120.14
Gln 658	8.21	4.06 (59.53)	2.19 (28.10)	C _ε H 6.58, 7.20 (109.57) C _v H 2.46 (34.19)
Glu 659	8.09	4.10 (59.52)	2.17 (29.60)	C _v H 2.39, 2.59 (36.00) NH 118.34
Leu 660	7.76	4.19 (58.21)	1.73, 1.85 (41.85)	C _v H 1.75 (27.20) C _{δ1} H ₃ 0.94 (23.04) C _{δ2} H ₃ 0.94 (24.00) NH 119.38
Leu 661	8.00	4.18 (58.17)	1.66, 1.96 (41.96)	C _v H 1.86 (27.16) C _{δ1} H ₃ 0.95 (22.88) C _{δ2} H ₃ 0.95 (23.84) NH 119.84
Glu 662	7.92	4.17 (59.14)	2.22, 2.27 (28.89)	C _v H 2.47, 2.56 (34.95) NH 7.92
Leu 663	8.19	4.31 (58.57)	1.90 (41.96)	C _v H 1.86 (27.27) C _{δ1} H ₃ 0.95 (24.19) C _{δ2} H ₃ 1.01 (23.81) NH 121.39
Asp 664	8.44	4.48 (57.28)	2.87, 2.95 (40.51)	NH 117.64
Lys 665	8.14	4.08 (59.48)	1.99, 2.06(31.96)	C _δ H 1.71 C _ε H 2.94, 2.90 (42.30) C _v H 1.46, 1.58 NH 120.59
Trp 666	8.35	4.42 (60.75)	3.89, 3.49	C _{δ1} H 7.16 N _{ε1} H 9.16, N _{ε1} 125.58 C _{ε3} H 7.68 C _{ζ2} H 7.24 C _{ζ3} H 7.14 C _η H 7.24 NH 121.37
Ala 667	8.98	3.84 (55.39)	1.55 (17.77)	NH 120.41
Ser 668	8.05	4.26 (61.46)	4.10, 4.09 (63.27)	NH 112.88
Leu 669	7.89	4.12 (57.29)	1.29, 1.63 (41.94)	C _v H 1.60 C _{δ1} H ₃ 0.75 (22.73) C _{δ2} H ₃ 0.83 (24.31) NH 123.42
Trp 670	7.78	4.50 (58.07)	2.93, 3.15	C _{δ1} H 6.76 C _{ε1} H 8.15 C _{ε3} H 7.53 C _{ζ2} H 7.29 C _{ζ3} H 7.11 C _{η2} H 7.18 NH 125.33
Asn 671	7.88	4.67	2.70, 2.76 (39.49)	N _{δ2} H 6.44, 7.41 (110.90) NH 118.01
CONH2	7.082 6.98			NH 105.64

TABLA A.2. Desplazamientos químicos de ^1H , ^{13}C y ^{15}N para el péptido **scrHybK3** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **25%** HFIP a 25 °C.

Residuo	HN	CaH	CβH	Otros
Gly 514		3.89, 3.83 (43.6)		
Ile 515	8.17	4.26 (62.24)	1.93 (39.29)	C _{v1} H 1.25, 1.54 (27.42) C _{v2} H ₃ 0.99 (16.97) C _δ H ₃ 0.92 (12.67)
Gly 516		3.90, 3.95 (45.78)		
Ala 517	7.80	4.19 (54.16)	1.35 (18.02)	
Phe 518	7.70	4.42 (59.93)	3.13, 3.19 (38.73)	C _δ H 7.24 C _ζ H 7.27 C _ε H 7.33 NH 115.81
Gly 519	7.96	3.78, 3.88		NH 106.37
Leu 520	7.57	4.31 (57.87)	1.68, 1.82 (42.40)	C _v H 1.68 C _{δ1} H ₃ 0.92 (22.71) C _{δ2} H ₃ 0.99 (23.07)
Leu 521	7.85	4.11 (58.55)	1.63, 1.76 (41.54)	C _v H 1.63 C _{δ1} H ₃ 0.90 (23.76) C _{δ2} H ₃ 0.95 (24.50)
Gly 522	7.78	3.75 (47.04)		
Phe 523	7.64	4.36 (61.08)	3.21, 3.31 (39.14)	C _δ H 7.20 C _ε H 7.25 NH 115.09
Leu 524	8.43	4.04 (57.99)	1.95, 2.00 (41.79)	C _v H 1.51 C _{δ1} H ₃ 0.93 (23.76) C _{δ2} H ₃ 0.93 (24.63) NH 121.29
Ala 525	8.69	4.04 (55.11)	1.48 (17.23)	NH 122.25
Ala 526	7.98	4.14 (54.89)	1.49 (17.90)	NH 120.43
Gly 527	8.14	3.78, 3.89 (46.43)		NH 105.50
Ser 528	8.01	4.29 (60.15)	3.96 4.01 (63.68)	NH 115.23
Lys 529	7.73	4.32 (56.87)	1.94 (32.33)	C _δ H 1.72 C _v H 1.49, 1.53 NH 120.87
Lys 530	7.87	4.28 (56.73)	1.84 (33.34)	C _v H 1.48 C _ε H 3.00 (42.37) NH 120.19
Ahx	7.53	2.31, 2.38 (38.13)	1.63, 1.67 (27.39)	C _δ H 1.54 (30.83) C _ε H 3.21 (42.18) C _v H 1.35 (28.52) NH 120.31
Lys 655	8.11	4.36 (56.88)	1.80, 1.88 (33.33)	C _v H 1.51 (24.75) NH 124.15
Asn 656	8.09	4.69 (53.12)	3.01 (39.27)	N _{δ2} H 6.74, 7.39 NH 115.62
Glu 657	9.23	4.05 (60.30)	2.12 (28.99)	C _v H 2.41, 2.46 (35.99)
Gln 658	8.19	4.06 (59.51)	2.19 (28.06)	C _{ε2} H 6.57, 7.24 (109.79) C _v H 2.46 (34.20) NH 117.73

Glu 659	8.09	4.10 (59.51)	2.16 (29.74)	C _v H 2.37, 2.59 (36.38)
Leu 660	7.77	4.19 (58.17)	1.71, 1.85 (41.84)	C _v H 1.75 (27.18) NH 119.31
Leu 661	7.99	4.18 (58.15)	1.65, 1.96 (41.86)	C _v H 1.85 (27.13) C _{δ1} H ₃ 0.94 (22.87) C _{δ2} H ₃ 0.94 (23.65) NH 118.88
Glu 662	7.91	4.18 (59.11)	2.24 (29.05)	C _v H 2.46, 2.54 (35.36)
Leu 663	8.17	4.29 (58.57)	1.90 (41.97)	C _v H 1.87 (27.27) C _{δ1} H ₃ 0.93 C _{δ2} H ₃ 1.00 NH 121.48
Asp 664	8.42	4.48 (57.42)	2.84, 2.91 (40.98)	
Lys 665	8.09	4.08 (59.43)	1.98, 2.07 (31.99)	C _δ H 1.71 C _ε H 2.90, 2.94 (42.30) C _v H 1.45, 1.58 (24.70) NH 120.41
Trp 666	8.33	4.42 (60.78)	3.38, 3.48 (29.16)	C _{δ1} H 7.17 N _{ε1} H 9.23, N _{ε1} 125.68 C _{ε3} H 7.68 C _{ζ2} H 7.24 C _{ζ2} H 7.13 C _η H 7.24 NH 121.36
Ala 667	8.97	3.84 (55.41)	1.55 (17.77)	NH 120.40
Ser 668	8.05	4.26 (61.46)	4.01, 4.09 (63.28)	NH 113.07
Leu 669	7.79	4.12 (57.31)	1.29, 1.63 (41.99)	C _v H 1.59 C _{δ1} H ₃ 0.82 (22.74) C _{δ2} H ₃ 0.76 (24.29) NH 123.39
Trp 670	7.80	4.50 (58.11)	2.93, 3.16 (29.54)	C _{δ1} H 6.77 C _{ε1} H 8.20, N _{ε1} 125.42 C _{ε3} H 7.54 C _{ζ2} H 7.29 C _{ζ3} H 7.11 C _{η2} H 7.18
Asn 671	7.89	4.67 (53.81)	2.70, 2.76 (39.49)	N _{δ2} H 6.44, 7.42 NH 117.94
CONH2	7.09 26.90			

TABLA A.3. Desplazamientos químicos de ^1H y ^{13}C para el péptido **HybK3** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y 2° mM DPC a 25 °C.

Residuo	HN	CaH	CβH	Otros
Gly 514		4.10		
Ile 515	9.53	3.91 (64.16)	1.90 (38.35)	C _{γ1} H 1.30, 1.58, (28.8) C _{γ2} H ₃ 0.95, (18.3) C _δ H ₃ 0.92, (13.98)
Gly 516	9.26	3.69, 3.86		
Ala 517	8.28	4.09	1.50 (18.63)	
Leu 518	7.88	4.11 (58.03)	1.87, 1.61 (42.35)	C _γ H 1.70 C _δ H ₃ 0.90
Phe 519	8.38	4.31 (61.29)	3.22 (39.50)	C _ζ H 7.08 C _δ H 7.17 C _ε H 7.14
Leu 520	8.48	3.95 (58.32)	1.55, 1.87 (42.47)	C _γ H 1.9 C _{δ1} H ₃ 0.89 C _{δ2} H ₃ 0.93
Gly 521	8.21	3.94		
Phe 522	7.98	4.41	3.21	C _δ H 7.18 C _ε H 7.14
Leu 523	7.98	3.88	1.47, 1.69 (42.28)	C _γ H 1.66 C _{δ1} H ₃ 0.75 (25.86) C _{δ2} H ₃ 0.78 (23.72)
Gly 524	8.06	3.86, 3.90		
Ala 525	7.87	4.26 (53.66)	1.42 (19.26)	
Ala 526	8.12	4.13 (53.87)	1.30 (19.01)	
Gly 527	8.25	3.87, 3.90		
Ser 528	7.98	4.36	3.86, 3.91	
Lys 529	8.22	4.28 (56.90)	1.79, 1.88	C _δ H 1.67 C _ε H 2.96 C _γ H 1.46
Lys 530	8.06	4.21 (56.84)	1.72, 1.79	C _δ H 1.67 C _γ H 1.41
Ahx	7.86	2.27, 2.31 (38.24)	1.58 (27.53)	C _δ H 1.48 (31.16) C _ε H 3.1, 3 3.18 (42.22) C _γ H 1.28 (28.70)
Lys 655	8.37	4.23 (57.20)	1.73, 1.78	C _δ H 1.69 C _ε H 2.96 C _γ H 1.42
Asn 656	8.32	4.68	2.89 (39.20)	N _{δ2} H 6.92, 7.63
Glu 657	8.66	4.13 (58.60)	2.02, 2.08	C _γ H 2.29
Gln 658	8.32	4.14	2.10 (28.92)	C _{ε21} H 6.85 C _{ε22} H 7.63 C _γ H 2.38

Glu 659	8.30	4.18	2.00, 2.06	C _v H 2.29 2.39
Leu 660	8.08	4.20	1.60, 1.79 (42.32)	C _v H 1.71 (27.39) C _{δ1} H ₃ 0.87 C _{δ2} H ₃ 0.91
Leu 661	7.89	4.32	1.61, 1.77 (42.56)	C _v H 1.70 C _{δ1} H ₃ 0.86 C _{δ2} H ₃ 0.89
Glu 662	7.98	4.30	2.05, 2.08	C _v H 2.29, 2.36
Leu 663	8.22	4.10 (58.02)	1.67, 1.79 (42.26)	C _v H 1.78 (27.42) C _{δ1} H ₃ 0.86 C _{δ2} H ₃ 0.93
Asp 664	8.50	4.39	2.72 (40.44)	
Lys 665	7.83	4.10 (58.60)	1.69, 1.80	C _δ H 1.55 C _ε H 2.80 (42.24) C _v H 1.17, 1.31
Trp 666	7.80	4.48	3.24, 3.43	C _δ H 7.33 N _{ε1} H 10.60 C _{ε3} H 7.52 C _ζ H 6.95 C _η H 7.07
Ala 667	8.21	3.87 (55.22)	1.46 (18.64)	
Ser 668	7.94	4.27	3.96 (63.28)	
Leu 669	7.74	4.15 (57.40)	1.79, 1.89 (42.53)	C _v H 1.55 C _{δ1} H ₃ 0.88 C _{δ2} H ₃ 0.94
Trp 670	7.72	4.53	3.23, 3.32	C _{δ1} H 7.18 C _{ε1} H 10.38 C _{ε3} H 7.51 C _{ζ2} H 7.44 C _{ζ3} H 6.98 C _{η2} H 7.07
Asn 671	7.98	4.53	2.71, 2.79 (39.64)	N _{δ2} H 6.97, 7.55
CONH2	7.10 7.25			

TABLA A.4. Desplazamientos químicos de ^1H para el péptido **scrHybK3** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **20 mM DPC** a 25 °C.

Residuos	HN	CaH	C β H	Otros
Gly 514				
Ile 515		4.17	1.90	C γ_1 H 1.26, 1.55 C γ_2 H ₃ 0.94 C δ H ₃ 0.86
Gly 516	8.58	3.99		
Ala 517	8.54	4.08	1.31	
Phe 518	8.52	4.37	3.17	C δ H 7.27
Gly 519	8.27	3.92		
Leu 520	8.18	4.18	1.69, 1.76	C δ H ₃ 0.88, 0.94
Leu 521	8.22	4.06	1.56, 1.78	C γ H 1.67 C δ H ₃ 0.86, 0.90
Gly 522	8.14	3.7, 3.86		
Phe 523	7.84	4.41	3.23	C δ H 7.27 C ϵ H 7.72
Leu 524	7.96	3.99	1.84	C γ H 1.48 C δ H ₃ 0.86
Ala 525	8.08	4.13	1.42	
Ala 526	7.81	4.21	1.44	
Gly 527	8.26	3.76, 3.86		
Ser 528	8.06	4.28	3.85, 3.96	
Lys 529	8.07	4.26	1.79, 1.88	C δ H 1.67 C ϵ H 1.48
Lys 530	8.01	4.20	1.74, 1.79	C δ H 1.67 C γ H 1.43 C ϵ H 2.95
Ahx	7.79	2.26, 2,31	1.58	C δ H 1.48 C ϵ H 3.09, 3.21 C γ H 1.28
Lys 655	8.39	4.24	1.73, 1.78	C δ H 1.67 C ϵ H 2.96 C γ H 1.41
Asn 656	8.32	4.68	2.89	N δ_2 H 6.92, 7.63
Glu 657	8.69	4.13	1.99, 2.07	C γ H 2.26
Gln 658	8.28	4.17	2.09	C ϵ_{21} H 6.85 C ϵ_{22} H 7.63 C γ H 2.37
Glu 659	8.32	4.19	1.97, 2.06	C γ H 2.25, 2.37
Leu 660	8.10	4.21	1.58, 1.77	C γ H 1.69 C δ_1 H ₃ 0.86 C δ_2 H ₃ 0.91
Leu 661	7.93	4.35	1.59, 1.75	C γ H 1.68 C δ_1 H ₃ 0.85 C δ_2 H ₃ 0.89
Glu 662	8.03	4.32	2.03, 2.10	C γ H 2.25, 2.33

Leu 663	8.23	4.09	1.65, 1.77	C _v H 1.76 C _{δ1} H ₃ 0.84 C _{δ2} H ₃ 0.92
Asp 664	8.51	4.38	2.69	
Lys 665		4.08	1.67, 1.79	C _δ H 1.54 C _ε H 2.77, 2.81 C _v H 1.15, 1.30
Trp 666	7.79	4.47	3.29, 3.41	C _{δ1} H 7.32 N _{ε1} H 10.64 C _{ε3} H 7.51 C _ζ H 7.44, 6.94 C _η H 7.07
Ala 667	8.22	3.87	1.46	
Ser 668	7.94	4.26	3.95	
Leu 669	7.73	4.41	1.79, 1.89	C _v H 1.55 C _{δ1} H ₃ 0.88 C _{δ2} H ₃ 0.94
Trp 670	7.72	4.53	3.23, 3.32	C _{δ1} H 7.18 C _{ε1} H 10.39 C _{ε3} H 7.50 C _{ζ2} H 7.44 C _{ζ3} H 6.98 C _{η2} H 7.07
Asn 671	7.98	4.53	2.71, 2.79	N _{δ2} H 6.97, 7.55
CONH2	7.10 7.25			

TABLA A.5. Desplazamientos químicos de ^1H para el péptido **MPERp** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **20 mM DPC** a 25 °C. Código de acceso BMRB-19263.

Residuo	HN	CaH	C β H	Otros
Asn 656		4.14	2.89, 2.89	N $_{\delta 2}$ H 6.99, 7.70
Glu 657		4.26	1.98, 2.08	C $_{\gamma}$ H 2.29, 2.29
Gln 658	8.41	4.25	2.03, 2.09	C $_{\gamma}$ H 2.36, 2.36 N $_{\epsilon}$ H $_2$ 6.85, 7.63
Glu 659	8.32	4.25	1.96, 2.04	C $_{\gamma}$ H 2.26, 2.26
Leu 660	8.29	4.27	1.58, 1.71	C $_{\gamma}$ H 1.67 C $_{\delta 1}$ H $_3$ 0.87 C $_{\delta 2}$ H $_3$ 0.92
Leu 661	8.06	4.36	1.56, 1.66	C $_{\gamma}$ H 1.61 C $_{\delta 1}$ H $_3$ 0.83 C $_{\delta 2}$ H $_3$ 0.87
Glu 662	8.25	4.32	1.98, 2.09	C $_{\gamma}$ H 2.23, 2.29
Leu 663	8.29	4.18	1.61, 1.71	C $_{\gamma}$ H 1.71 C $_{\delta}$ H $_3$ 0.84, 0.92
Asp 664	8.48	4.50	2.72, 2.72	
Lys 665	8.08	4.08	1.63, 1.72	C $_{\gamma}$ H 1.16, 1.16 C $_{\delta}$ H 1.51, 1.51 C $_{\epsilon}$ H 2.80, 2.80
Trp 666	7.94	4.53	3.24, 3.43	C $_{\delta 1}$ H 7.38 N $_{\epsilon 1}$ H 10.65 C $_{\epsilon 3}$ H 7.51 C $_{\zeta 3}$ H 6.92 C $_{n 2}$ H 7.06 C $_{\zeta 2}$ H 7.42
Ala 667	8.11	4.08	1.51	
Ser 668	8.10	4.31	3.90, 3.98	
Leu 669	7.83	4.15	1.65, 1.79	C $_{\gamma}$ H 1.66 C $_{\delta}$ H $_3$ 0.79, 0.85
Trp 670	7.94	4.59	3.24, 3.43	C $_{\delta 1}$ H 7.25 N $_{\epsilon 1}$ H 10.30 C $_{\epsilon 3}$ H 7.54 C $_{\zeta 3}$ H 6.91 C $_{n 2}$ H 7.04 C $_{\zeta 2}$ H 7.41
Asn 671	8.02	4.58	2.82, 2.88	N $_{\delta 2}$ H 7.00, 7.63
Trp 672	nd	4.37	3.31, 3.31	C $_{\delta 1}$ H 7.18 N $_{\epsilon 1}$ H 10.50 C $_{\epsilon 3}$ H 7.36 C $_{\zeta 3}$ H 6.91 C $_{n 2}$ H 7.06 C $_{\zeta 2}$ H 7.44
Phe 673	7.98	4.22	2.95, 3.10	C $_{\delta}$ H 7.21, 7.21 C $_{\epsilon}$ H 7.12, 7.12 C $_{\zeta}$ H nd
Asn 674	7.64	4.55	2.62, 2.81	N $_{\delta 2}$ H 6.70, 7.67

Ile 675	8.12	3.67	1.61	C _γ H 1.14, 1.42 C _γ H ₃ 0.76 C _δ H ₃ 0.75
Thr 676	8.03	3.85	4.02	C _γ H ₃ 1.05
Asn 677	7.90	4.57	2.69, 2.69	N _{δ2} H 6.70, 7.59
Trp 678	7.99	4.58	3.33, 3.33	C _{δ1} H 7.19 N _{ε1} H 10.27 C _{ε3} H 7.42 C _{ζ3} H 6.85 C _{η2} H 6.98 C _{ζ2} H 7.33
Leu 679	8.23	4.03	1.50, 1.79	C _γ H 1.79 C _δ H ₃ 0.85, 3 0.91
Trp 680	7.93	4.26	3.33, 3.33	C _{δ1} H 7.16 N _{ε1} H 10.24 C _{ε3} H 7.45 C _{ζ3} H 7.06 C _{η2} H 7.21 C _{ζ2} H 7.49
Tyr 681	7.74	4.00	3.00, 3.12	C _δ H 7.09, 7.09 C _ε H 6.81, 6.81
Ile 682	7.83	3.77	1.92	C _γ H 1.30, 1.64 C _γ H ₃ 0.90 C _δ H ₃ 0.88
Lys 683	7.69	4.04	1.73, 1.80	C _γ H 1.41, 1.47 C _δ H 1.58, 1.58 C _ε H 2.81, 2.81
CONH2	6.96, 7.26			

TABLA A.6. Desplazamientos químicos de ^1H para el péptido **MPERp** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **25% HFIP** a 25 °C. Código de acceso BMRB-19262.

Residuo	HN	CaH	C β H	Otros
Asn 656		nd	3.21, 3.36	N δ_2 H 6.99, 7.72
Glu 657		4.17	2.15, 2.15	C γ H 2.49, 2.49
Gln 658	8.34	4.04	2.15, 2.18	C γ H 2.46, 2.46, N ϵ H $_2$ nd
Glu 659	8.09	4.09	2.13, 2.20	C γ H 2.45, 2.52
Leu 660	7.56	4.22	1.77, 1.83	C γ H 1.76 C δ_1 H $_3$ 0.93 (23.0) C δ_2 H $_3$ 0.99 (23.7)
Leu 661	7.91	4.18	1.65, 1.93	C γ H 1.83 C δ_1 H $_3$ 0.84 (23.5) C δ_2 H $_3$ 0.97 (24.3)
Glu 662	7.92	4.17	2.21, 2.27	C γ H 2.49, 2.58
Leu 663	8.17	4.30	1.88, 1.88	C γ H 1.83 C δ_1 H $_3$ 0.94 (22.6) C δ_2 H $_3$ 1.00 (23.4)
Asp 664	8.48	4.50	2.92, 3.01	
Lys 665	8.17	4.09	2.00, 2.07	C γ H 1.42, 1.60 C δ H 1.68, 1.68 C ϵ H 2.88, 2.88 (42.1)
Trp 666	8.33	4.48	3.40, 3.51	C δ_1 H 7.11 N ϵ_1 H 9.06 C ϵ_3 H 7.64 C ζ_3 H 7.04 C η_2 H 6.97 C ζ_2 H 7.11
Ala 667	9.12	3.98	1.61, 17.6	
Ser 668	8.15	4.32	nd	
Leu 669	7.99	4.29	1.73, 1.83	C γ H 1.71 C δ_1 H $_3$ 0.88 (23.3) C δ_2 H $_3$ 0.92 (23.8)
Trp 670	8.61	4.39	2.80, 3.13	C δ_1 H 7.25 N ϵ_1 H 8.96 C ϵ_3 H 7.52 C ζ_3 H 7.06 C η_2 H 7.24 C ζ_2 H 7.44
Asn 671	8.46	4.42	2.77, 3.08	N δ_2 H 6.34, 7.37
Trp 672	8.30	4.36	3.42, 3.59	C δ_1 H 7.01 N ϵ_1 H 9.02 C ϵ_3 H 7.57 C ζ_3 H 7.06 C η_2 H 7.18 C ζ_2 H 7.36
Phe 673	9.14	3.84	3.12, 3.21	C δ H 7.09, 7.09 C ϵ H 7.29, 7.29 C ζ H nd

Asn 674	7.98	4.20	2.39, 2.66	N _{δ2} H nd
Ile 675	8.45	3.83	1.82	C _v H 1.14, 1.70 C _v H ₃ 0.93, 16.6 C _δ H ₃ 0.85, 12.4
Thr 676	7.98	3.73	3.71	C _{v2} H ₃ 0.70, 20.4
Asn 677	7.62	4.44	2.47, 2.65	N _{δ2} H 5.32, 6.73
Trp 678	7.99		3.34, 3.53	C _{δ1} H 7.09 N _{ε1} H 9.09 C _{ε3} H 7.57 C _{ζ3} H 7.01 C _{η2} H 7.17 C _{ζ2} H 7.36
Leu 679	8.46	3.87	1.40, 1.84	C _v H 1.85 C _{δ1} H ₃ 0.84 (22.4) C _{δ2} H ₃ 0.86 (24.4)
Trp 680	8.11	4.20	3.27, 3.32	C _{δ1} H 6.96 N _{ε1} H 9.41 C _{ε3} H 7.49 C _{ζ3} H 7.11 C _{η2} H 7.22 C _{ζ2} H 7.42
Tyr 681	7.83	4.09	3.00, 3.11	C _δ H 7.06, 7.06 C _ε H 6.85, 6.85
Ile 682	8.11	3.73	1.70	C _v H 1.10, 1.23 C _v H ₃ 0.66 (17.0) C _δ H ₃ 0.65 (11.8)
Lys 683	8.00	4.03	1.78, 1.78	C _v H 1.39, 1.39 C _δ H 1.58, 1.58 C _ε H 2.84, 2.84 (42.0)
CONH2	6.66, 6.92			

TABLA A.7. Desplazamientos químicos de ^1H para el péptido **CpreTM** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **20 mM DPC** a 25 °C. Código de acceso BMRB-19583.

Residuo	HN	CaH	C β H	Otros
Lys 668		3.75	1.76, 1.76	C γ H 1.41, 1.41 C δ H 1.68, 1.68 C ϵ H 2.98, 2.98
Lys 669		4.30	1.75, 1.79	C γ H 1.42, 1.42 C δ H 1.67, 1.67 C ϵ H 2.98, 2.98
Lys 670	8.19	4.27	1.67, 1.67	C γ H 1.33, 1.38 C δ H 1.61, 1.61 C ϵ H 2.91, 2.91
Asn 671	8.56	4.74	2.62, 2.62	N δ_2 H 6.96, 7.72
Trp 672	8.54	4.40	3.19, 3.25	C δ_1 H 7.38 N ϵ_1 H 10.74 C ϵ_3 H 7.50 C ζ_3 H 6.99 C η_2 H 7.09 C ζ_2 H 7.50
Phe 673	7.48	4.41	2.69, 2.74	C δ H 6.91, 6.91 C ϵ H 7.14, 7.14 C ζ H 7.08
Asp 674	7.82	4.64	2.54, 2.82	
Ile 675	8.19	3.94	1.93	C γ H 1.26, 1.46 C γ H $_3$ 0.86 C δ H $_3$ 0.88
Thr 676	8.21	3.94	4.16	C γ_2 H $_3$ 1.23
Asn 677	8.28	4.66	2.71, 2.77	N δ_2 H 6.91, 7.76
Trp 678	8.04	4.74	3.33, 3.46	C δ_1 H 7.33 N ϵ_1 H 10.62 C ϵ_3 H 7.43 C ζ_3 H 6.85 C η_2 H 7.04 C ζ_2 H 7.46
Leu 679	8.13	4.07	1.67, 1.90	C γ H 1.84 C δ_1 H $_3$ 0.93 C δ_2 H $_3$ 1.00
Trp 680	8.27	4.06	3.14, 3.31	C δ_1 H 7.25 N ϵ_1 H 10.55 C ϵ_3 H 6.46 C ζ_3 H 6.74 C η_2 H 7.00 C ζ_2 H 7.43
Tyr 681	7.62	4.05	3.08, 3.25	C δ H 7.24, 7.24 C ϵ H 6.88, 6.88
Ile 682	8.22	3.77	2.13	C γ H 1.19, 2.02 C γ H $_3$ 1.01 C δ H $_3$ 0.98
Lys 683	8.43	3.83	1.82, 1.82	C γ H 1.31, 1.57 C δ H 1.68, 1.68

				C _ε H 2.83, 2.89
Leu 684	7.73	3.87	1.08, 1.59	C _γ H 1.37 C _{δ1} H ₃ 0.59 C _{δ2} H ₃ 0.73
Phe 685	8.37	4.06	3.20, 3.41	C _δ H 7.06, 7.06 C _ε H 7.03, 7.03 C _ζ H 6.96
Ile 686	8.48	3.45	2.02	C _γ H 1.19, 2.00 C _γ H ₃ 0.91 C _δ H ₃ 0.88
Met 687	8.14	4.06	2.14, 2.27	C _γ H 2.46, 2.73 C _ε H ₃ 2.03
Ile 688	8.34	3.67	1.88	C _γ H 1.03, nd C _γ H ₃ 0.80 C _δ H ₃ 0.81
Val 689	8.40	3.40	1.77	C _{γ1} H ₃ 0.48 C _{γ2} H ₃ 0.72
Gly 690	7.99	3.63, 3.88		
Gly 691	7.86	3.71, 3.89		
Leu 692	8.00	4.11	1.43, 1.98	C _γ H 1.90 C _{δ1} H ₃ 0.83 C _{δ2} H ₃ 0.83
Val 693	8.19	3.72	2.19	C _{γ1} H ₃ 0.93 C _{γ2} H ₃ 1.01
Lys 694	7.73	4.08	1.91, 1.91	C _γ H 1.49, 1.58 C _δ H 1.68, 1.68 C _ε H 2.94, 2.94
Lys 695	7.71	4.15	1.91, 1.91	C _γ H 1.53, 1.53 C _δ H 1.69, 1.69 C _ε H 2.94, 2.94
CONH2	6.96, 7.26			

TABLA A.8. Desplazamientos químicos de ^1H y ^{13}C para el péptido **CpreTM** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **25% HFIP** a 25 °C. Código de acceso BMRB-19582.

Residuo	HN	CaH	CβH	Otros
Lys 668		4.01 (56.0)	1.95, 1.95 (33.9)	C _v H 1.53, 1.53 (24.1) C _δ H 1.76, 1.76 C _ε H 3.06, 3.06 (42.1)
Lys 669		4.43 (56.6)	1.84, 1.92 (33.5)	C _v H 1.54, 1.54 (24.9) C _δ H 1.77, 1.77 C _ε H 3.06, 3.06 (42.1)
Lys 670	8.38	4.38 (56.7)	1.78, 1.78 (33.3)	C _v H 1.45, 1.50 (24.8) C _δ H 1.68, 1.68 (29.1) C _ε H 2.99, 2.99
Asn 671	8.57	Nd	2.69, 2.76	N _{δ2} H 6.79, 7.53
Trp 672	8.00	4.46 (59.2)	3.25, 3.35 (29.2)	C _{δ1} H 7.17 N _{ε1} H 9.59 C _{ε3} H 7.49 C _{ζ3} H 7.17 C _{n2} H 7.27 C _{ζ2} H 7.49
Phe 673	7.40	4.28 (59.0)	2.80, 2.85 (38.5)	C _δ H 7.06, 7.06 C _ε H 7.31, 7.31 C _ζ H 7.25
Asp 674	7.86	4.59 (55.7)	2.80, 2.80 (41.4)	
Ile 675	7.84	4.05 (65.4)	1.97 (38.6)	C _v H 1.34, 1.62 (28.5) C _v H ₃ 1.03, 16.8 C _δ H ₃ 0.94 (12.5)
Thr 676	8.08	4.05 (64.0)	4.22 (69.1)	C _{v2} H ₃ 1.22 (21.2)
Asn 677	7.93	4.65	2.71, 2.78 (38.7)	N _{δ2} H 6.32, 7.17
Trp 678	7.91	4.75	3.35, 3.58 (29.5)	C _{δ1} H 7.23 N _{ε1} H 9.28 C _{ε3} H 7.59 C _{ζ3} H 7.05 C _{n2} H 7.20 C _{ζ2} H 7.39
Leu 679	8.00	4.05 (57.9)	1.52, 1.82	C _v H 1.83 (26.8) C _{δ1} H ₃ 0.89 (23.1) C _{δ2} H ₃ 0.95 (24.3)
Trp 680	7.85	4.25 (60.3)	3.33, 3.41 (28.7)	C _{δ1} H 7.04 N _{ε1} H 9.27 C _{ε3} H 7.35 C _{ζ3} H 7.09 C _{n2} H 7.25 C _{ζ2} H 7.43
Tyr 681	7.40	4.05 (61.3)	2.79, 3.03	C _δ H 7.12, 7.12 C _ε H 6.84, 6.84
Ile 682	7.93	3.80 (64.3)	1.86 (37.7)	C _v H 1.15, 1.51 (28.2) C _v H ₃ 0.92 (16.7) C _δ H ₃ 0.79 (11.6)

Lys 683	7.88	3.92 (60.1)	1.87, 1.87 (32.0)	C _v H 1.39, 1.56 (25.2) C _δ H 1.65, 1.65 (29.4) C _ε H 2.92, 2.92 (42.3)
Leu 684	7.58	4.08 (58.2)	1.53, 1.74 (41.7)	C _v H 1.62, 26.9 C _{δ1} H ₃ 0.90 (22.3) C _{δ2} H ₃ 0.91 (23.9)
Phe 685	8.14	4.14 (61.9)	3.10, 3.30 (38.9)	C _δ H 7.16, 7.16 C _ε H 7.25, 7.25 C _ζ H 7.21
Ile 686	8.47	3.59 (65.1)	2.00 (37.9)	C _v H 1.31, 1.92 (28.8) C _v H ₃ 0.94 (16.6) C _δ H ₃ 0.89 (12.2)
Met 687	8.11	4.13 (59.1)	2.21, 2.41 (32.6)	C _v H 2.56, 2.80 (32.0) C _ε H ₃ 2.07 (15.8)
Ile 688	8.55	3.81 (65.1)	1.99 (38.3)	C _v H 1.14, 1.81 (28.5) C _v H ₃ 0.88 (15.9) C _δ H ₃ 0.86 (11.9)
Val 689	8.64	3.60 (66.9)	1.74 (31.2)	C _{v1} H ₃ 0.66 (21.8) C _{v2} H ₃ 0.79 (20.5)
Gly 690	8.32	3.80, 3.90 (46.9)		
Gly 691	7.62	3.94, 3.94 (46.6)		
Leu 692	8.05	4.23	1.61, 2.03 (42.1)	C _v H 1.84 (26.7) C _{δ1} H ₃ 0.89 (23.1) C _{δ2} H ₃ 0.93 (24.2)
Val 693	8.31	3.83 (65.3)	2.23 (32.0)	C _{v1} H ₃ 0.99 (20.5) C _{v2} H ₃ 1.07 (21.1)
Lys 694	7.85	4.22 (57.7)	1.98, 1.98 (32.5)	C _v H 1.55, 1.55 C _δ H 1.76, 1.76 C _ε H 3.04, 3.04
Lys 695	7.88	4.23 (57.1)	1.99, 1.99 (32.7)	C _v H 1.60, 1.60 C _ε H 3.06, 3.06
CONH2	6.89, 7.36			

TABLA A.9. Desplazamientos químicos de ^1H y ^{13}C para el péptido **TMDp** en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía tampón HEPES 2 mM a pH 7.0 y **25% HFIP** a 25 °C. Código de acceso BMRB-19581.

Residuo	HN	CaH	C β H	Otros
Lys 681		3.60 (57.0)	1.67, 1.94 (35.9)	C γ H 1.53, 1.58 (25.1) C δ H 1.76, 1.76 (29.4) C ϵ H 3.06, 3.06 (42.2)
Lys 682		4.21 (58.8)	1.90, 1.90 (33.1)	C δ H 1.77, 1.77 C ϵ H 3.03, 3.03
Lys 683	8.17	4.09 (59.1)	1.82, 1.82 (32.7)	C γ H 1.49, 1.56 (25.3) C δ H 1.76, 1.76 (29.1) C ϵ H 3.03, 3.03 (42.1)
Leu 684	7.21	4.29 (57.4)	1.65, 1.75 (42.3)	C γ H 1.66 (27.1) C δ H ₃ 0.99 (23.0) C δ 2H ₃ 0.92 (24.3)
Phe 685	7.67	4.29 (61.3)	3.20, 3.36 (38.9)	C δ H 7.18, 7.18 C ϵ H 7.28, 7.28 C ζ H 7.24
Ile 686	7.79	3.72 (64.4)	1.97 (38.0)	C γ H 1.34, 1.74 (28.5) C γ H ₃ 0.97 (16.7) C δ H ₃ 0.92 (12.1)
Met 687	7.80	4.17 (58.7)	2.23, 2.37 (32.5)	C γ H 2.59, 2.78 (31.9) C ϵ H ₃ 2.09 (15.8)
Ile 688	8.19	3.83 (65.1)	1.98 (38.4)	C γ H 1.15, 1.80 (28.5) C γ H ₃ 0.91 (15.9) C δ H ₃ 0.88 (12.1)
Val 689	8.35	3.64 (67.0)	1.89 (31.5)	C γ H ₃ 0.85 (20.4) C γ H ₃ 0.77 (21.8)
Gly 690	8.37	3.83, 3.90 (46.9)		
Gly 691	7.80	3.94, 4.01 (46.3)		
Leu 692	8.07	4.26 (57.8)	1.59, 2.04 (42.1)	C γ H 1.87 (26.7) C δ H ₃ 0.88 (22.0) C δ H ₃ 0.91 (24.2)
Val 693	8.45	3.70 (66.7)	2.19 (31.9)	C γ H ₃ 0.99 (20.4) C γ H ₃ 1.07 (21.8)
Gly 694	7.40	3.88, 3.88 (47.0)		
Leu 695	7.80	4.12 (57.7)	1.63, 1.83 (42.0)	C γ H 1.73 (26.7) C δ 1H ₃ 0.88 (22.7) C δ 2H ₃ 0.92 (23.7)
Arg 696	7.75	4.11 (59.2)	2.09, 2.14 (29.4)	C γ H 1.81, 1.93 (27.3) C δ H 3.21, 3.26 (43.2) N ϵ H 7.10
Ile 697	7.91	3.87 (65.0)	2.14 (37.9)	C γ H 1.18, 1.80 (28.5) C γ H ₃ 0.97 (16.3) C δ H ₃ 0.89
Val 698	8.03	3.56 (67.5)	2.19 (31.6)	C γ 1H ₃ 0.90 (20.2) C γ 2H ₃ 1.07 (22.0)
Phe 699	8.37	4.34 (61.0)	3.24, 3.24 (38.4)	C δ H 7.16, 7.16 C ϵ H 7.25, 7.25

				C _z H ---
Ala 700	8.19	4.15 (55.9)	1.69 (17.4)	
Val 701	8.69	3.63 (67.3)	2.29 (31.7)	C _{v1} H ₃ 0.96 (20.5) C _{v2} H ₃ 1.12 (21.8)
Leu 702	8.91	4.05 (58.4)	1.46, 1.94 (41.7)	C _v H 1.94 (26.6) C _{δ1} H ₃ 0.86 (21.8) C _{δ2} H ₃ 0.85 (24.5)
Ser 703	8.07	4.13 (62.5)	3.86, 4.16 (62.7)	
Ile 704	8.10	3.83 (64.9)	2.09 (38.3)	C _v H 1.16, 1.85 (28.8) C _v H ₃ 0.97 C _δ H ₃ 0.90
Lys 705	8.54	4.12 (58.6)	1.96, 2.05 (32.0)	C _v H 1.58, 1.63 (24.8) C _δ H 1.71, 1.71 (28.8) C _ε H 2.96, 2.96 (42.0)
Lys 706	8.35	4.20 (57.8)	1.99, 1.99 (32.4)	C _v H 1.58, 1.58 (24.8) C _δ H 1.75, 1.75
Lys 707	7.85	4.21 57.3	2.01, 2.01 (32.7)	C _v H 1.58, 1.65 C _δ H 1.76, 1.76 (29.0) C _ε H 3.06, 3.06
CONH2	6.99, 7.37			

TABLA A.10. Desplazamientos químicos de ^1H , ^{13}C y ^{15}N del zimógeno **ONCFLG** marcado con ^{13}C y ^{15}N en 0.5 mL de $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1 relación v/v). La solución contenía 50 μM DSS, 3 mM NaN_3 , 100 mM KCl, 1 mM HAc, 1 mM K_2HPO_4 a pH 5.20 y a 35 °C. Código de acceso BMRB-17973.

Residuo	HN	CaH	C β H	Otros
Arg 1				
Pro 2		4.39 (63.31)	1.88, 2.27	C γ H 2.00, 2.06 (27.44) C δ H 3.67, 3.83 (50.64) CO 176.27
Cys 3	8.46	4.46 (57.36)	2.88, 3.72 (46.46)	NH 120.85 CO 172.71
Lys 4	7.34	4.78 (54.12)	1.51, 1.75 (34.55)	C γ H 1.27, 1.35 (24.39) C δ H 1.65, 1.65 (28.49) C ϵ H 2.91, 2.96 (42.27) NH 119.32 CO 175.82
Try 5	9.32	4.83 (59.04)	2.55, 2.66 (43.84)	C β H 2.55, 2.66 C δ H 6.71, 6.71 NH 119.65 CO 176.52
Lys 6	9.01	4.75 (54.90)	1.82, 1.82 (35.34)	C δ H 1.75, 1.75 (29.14) C ϵ H 3.02, 3.02 (42.21) C γ H 1.46, 1.46 (24.94) NH 120.03 CO 174.84
Leu 7	8.41	5.38 (54.02)	1.66, 1.83 (43.83)	C γ H 1.42 (28.09) C δ_1 H ₃ 1.00 (23.99) C δ_2 H ₃ 1.09 (26.95) NH 126.61 CO 176.78
Lys 8	9.34	4.86 (55.18)	1.81, 1.91 (35.49)	C δ H 1.48, 1.48 (25.01) C ϵ H 3.02, 3.02 (42.14) C γ H 1.35, 1.35 (24.59) NH 128.51 CO 175.95
Lys 9	9.19	5.28 (56.22)	1.84, 1.84 (34.27)	C δ H 1.45, 1.59 (29.74) C ϵ H 2.63, 2.86 (42.28) C γ H 1.33, 1.33 (26.02) NH 129.76 CO 175.70
Ser 10	8.89	4.85 (58.68)	3.93, 3.93 (66.30)	NH 117.07 CO 172.50
Thr 11	8.42	5.49 (60.22)	4.00 (70.39)	C γ_2 H ₃ 0.99 (22.38) NH 115.17 CO 174.23
Asn 12	9.10	4.96 (52.33)	2.68, 3.09 (43.15)	N δ_2 H 7.24, 7.79 (111.72) NH 121.30 CO 176.01

Lys 13	9.19	4.54 (56.41)	1.45, 1.93 (35.31)	C _δ H 1.61, 1.61 (29.47) C _ε H 2.82, 2.87 (42.32) C _γ H 1.02, 1.02 (26.09) NH 120.75 CO 176.82
Phe 14	8.04	4.97 (56.57)	3.00, 3.15 (41.43)	C _{δ1} H 6.85, 6.85 (132.41) C _{ε1} H 6.62, 6.62 (130.89) C _ζ H 6.45 (129.89) NH 114.03 CO 170.62
Cys 15	9.01	5.77 (54.08)	2.12, 3.25 (49.59)	C _β H 2.12, 3.25 NH 120.16 CO 173.64
Val 16	8.86	5.06 (58.27)	2.10 (35.71)	C _{γ1} H ₃ 0.74 (21.39) C _{γ2} H ₃ 1.08 (23.53) NH 121.40 CO 173.46
Thr 17	8.65	5.00 (62.50)	3.83 (69.61)	C _{γ2} H ₃ 1.08 (23.38) NH 118.38 CO 174.78
Cys 18	9.24	5.17 (52.69)	2.59, 2.77 (38.87)	NH 127.18 CO 173.39
Glu 19	8.78	4.18 (56.93)	1.89, 1.96 (33.27)	C _γ H 2.18, 2.29 (36.40) NH 118.77 CO 176.07
Asn 20	10.27	4.35 (54.83)	2.75, 3.10 (37.22)	N _{δ2} H 6.88, 7.67 (113.04) NH 126.57 CO 174.37
Gln 21	8.86	4.05 (57.20)	1.68, 2.05 (27.38)	C _γ H 2.06, 2.24 (36.03) N _{ε2} H 6.98, 7.74 (113.81) NH 125.89 CO 174.08
Ala 22	7.63	4.42 (49.59)	0.93 (20.99)	NH 119.15
Pro 23		4.58 (62.53)	0.64, 1.23 (30.38)	C _γ H 1.63, 1.67 (27.99), C _δ H 2.98, 3.21 (50.76), CO 175.99
Val 24	8.56	4.67 (60.59)	2.11 (33.70)	C _{γ1} H ₃ 0.70 (18.34) C _{γ2} H ₃ 0.99 (22.41) NH 129.65 CO 176.78
His 25	7.39	5.56 (54.93)	3.03, 3.27 (32.46)	C _{δ2} H 7.12 (120.45) C _{ε1} H 8.65 (137.81) NH 117.22
Phe 26		4.39 (58.77)	1.62, 2.88 (39.13)	C _δ H 6.63, 6.63 (130.85) C _{ε1} H 6.70, 6.70 (130.63) CO 173.61
Val 27	8.17	3.78 (64.62)	1.20 (32.81)	C _{γ1} H ₃ 0.73 (22.51) C _{γ2} H ₃ 0.85 (21.01) NH 129.18 CO 176.07
Gly 28	6.65	4.12, 4.12 (45.56)		NH 123.40 CO 170.86
Val 29	8.74	4.29 (62.89)	2.14 (33.26)	C _{γ1} H ₃ 1.10 (22.00)

				C _{v2} H ₃ 1.23 (22.40) NH 118.62 CO 177.83
Gly 30	8.75	3.36, 3.36 (46.80)		NH 115.29 CO 172.77
Ser 31	7.36	4.47 (57.49)	3.69, 3.97 (64.89)	NH 111.54 CO 173.40
Cys 32	9.11	4.57 (54.60)	2.84, 3.20 (41.87)	NH 118.07 CO 174.84
Gly 33	8.33	4.67, 4.67(45.30)		NH 129.17 CO 174.05
Ser 34	8.34	4.47 (58.57)	3.85, 3.92 (64.14)	NH 115.88 CO 175.30
Gly 35	8.46	3.97, 3.97 (45.58)		NH 110.79 CO 174.78
Gly 36	8.20	4.00, 4.00 (45.38)		NH 129.38 CO 174.36
Ser 37	8.25	4.45 (58.65)	3.70, 3.78 (64.17)	NH 115.51 CO 175.10
Gly 38	8.39	4.02, 4.02 (45.51)		NH 110.63 CO 173.96
Ile 39	7.72	4.12 (61.25)	1.75 (38.62)	C _δ H ₃ 0.79 (13.02) C _{v1} H 1.05, 1.27 (27.15) C _{v2} H 0.77 (17.48) NH 119.39 CO 175.79
Phe 40	8.22	4.66 (57.69)	2.99, 3.14 (39.46)	C _δ H 7.25, 7.25 (131.64) C _ε H 7.34, 7.34 (131.84) NH 124.02 CO 175.36
Leu 41	8.03	4.36 (55.37)	1.55, 1.55 (62.64)	C _v H 1.35 (24.88) C _{δ1} H ₃ 0.82 (23.51) C _{δ2} H ₃ 0.90 (25.11) NH 124.18 CO 177.04
Glu 42	8.34	4.34 (56.97)	1.98, 2.09 (29.91)	C _v H 2.32, 2.32 (36.12) NH 121.78 CO 176.80
Thr 43	8.04	4.34 (62.23)	4.28 (69.77)	C _{v2} H ₃ 1.21 (21.74) NH 113.92 CO 174.89
Ser 44	8.26	4.47 (58.76)	3.90, 3.90 (63.78)	NH 117.70 CO 174.68
Leu 45	8.14	4.34 (55.48)	1.67, 1.67 (42.17)	C _v H 1.60 (29.83) C _δ H ₃ 0.91, 0.91 (25.07) NH 123.51 CO 177.56
Ser 46	8.17	4.46 (58.36)	3.86, 3.86 (64.06)	NH 116.00 CO 174.58
Ala 47	8.26	4.37 (53.08)	1.43 (19.09)	NH 125.97 CO 178.36
Gly 48	8.51	3.99, 4.08 (45.75)		129.30 CO 174.57

Ser 49	8.09	4.46 (57.74)	3.92, 3.92 (63.25)	NH 116.55 CO 173.12
Asp 50	8.16	4.88 (53.27)	2.94, 3.13 (42.87)	NH 120.60 CO 175.69
Trp 51	8.99	5.49 (58.19)	3.22, 3.39 (31.48)	C _{δ1} H 7.32 (128.33) C _{ε3} H 7.24 (123.94) C _{ε1} H 10.27 C _η H 6.88 (120.68) C _{ζ1} H 7.40 (115.19) C _{ζ2} H 8.28 (125.04) NH 123.03 CO 178.10
Leu 52	8.14	3.79 (58.70)	1.62, 1.80 (40.87)	C _γ H 1.66 (27.28) C _{δ1} H ₃ 0.86 (24.02) C _{δ2} H ₃ 0.93 (24.39) NH 120.92 CO 179.96
Thr 53	8.67	4.34 (67.39)	3.74 (67.90)	C _{γ2} H ₃ 1.28 (22.19) NH 117.31 CO 175.86
Phe 54	8.65	4.11 (62.49)	3.20, 3.69 (39.65)	C _δ H 7.05, 7.05 (132.96) C _{ε1} H 7.10, 7.10 (130.19) C _ζ H 6.93 (128.33) NH 123.50 CO 178.05
Gln 55	8.26	3.57 (59.69)	1.71, 1.82 (30.15)	C _γ H 1.28, 1.28 (35.02) N _{ε2} H 4.67, 6.57 (128.18) NH 117.36 CO 178.39
Lys 56	7.55	3.98 (59.14)	1.75, 1.85 (32.75)	C _γ H 1.46, 1.46 (25.29) C _δ H 1.56, 1.56 (29.43) C _ε H 2.88, 2.88 (42.19) NH 117.92 CO 178.77
Lys 57	8.00	4.13 (57.83)	1.01, 1.34 (34.63)	C _γ H 1.41, 1.41 (26.90) C _δ H 1.52, 1.52 (29.49) C _ε H 2.79, 2.79 (39.73) NH 114.03 CO 176.56
His 58	7.85	4.65 (55.47)	1.80, 2.64 (30.30)	C _{δ2} H 6.38 (122.70) C _{ε1} H 8.33 (138.54) NH 128.05 CO 173.94
Ile 59	8.29	5.15 (60.89)	2.26 (37.33)	C _δ H ₃ 1.11 (12.97) C _{γ1} H 1.35, 1.67 (28.25) C _{γ2} H 1.07 (17.73) NH 121.71 CO 175.09
Thr 60	8.65	5.08 (58.11)	3.76 (70.94)	C _{γ2} H ₃ 1.10 (20.06) NH 120.51 CO 171.44
Asn 61	8.33	4.97 (52.77)	2.86, 3.07 (38.39)	N _{δ2} H 6.81, 7.53 NH 123.15

				CO 174.85
Thr 62	7.96	4.49 (58.98)	3.98 (69.61)	C _{v2} H ₃ 0.84 (23.85) NH 113.55 CO 172.66
Arg 63	8.50	3.59 (57.91)	0.98, 1.39 (29.42)	C _v H 0.99, 1.09 (26.58) C _δ H 2.77, 2.77 (42.98) NH 125.70 CO 178.58
Asp 64	7.92	4.65 (53.14)	2.24, 2.42 (39.26)	NH 122.08 CO 174.10
Val 65	6.57	3.12 (62.97)	1.34 (32.52)	C _{v1} H ₃ -0.02 (20.19) C _{v2} H ₃ 0.45 (21.26) NH 121.27 CO 175.01
Asp 66	8.58	4.63 (51.47)	2.47, 2.88 (38.13)	NH 126.84 CO 175.44
Cys 67	7.94	4.65 (57.85)	2.98, 3.13 (39.82)	NH 121.13 CO 175.01
Asp 68	8.74	4.39 (58.11)	2.64, 2.78 (39.39)	NH 118.59 CO 177.94
Asn 69	7.10	4.66 (55.07)	2.61, 2.61 (39.66)	N _{δ2} H 6.81, 8.31 (115.11) NH 114.59 CO 176.82
Ile 70	7.90	4.08 (63.52)	1.56 (38.40)	C _δ H ₃ 0.59 (17.93), C _{v1} H 1.21, 1.40 (29.07) C _{v2} H 0.72 (13.03) NH 119.13 CO 178.38
Met 71	8.51	4.26 (55.57)	1.77, 1.77 (29.78)	NH 117.99 CO 176.49
Ser 72	7.22	4.79 (58.30)	4.17, 4.22 (64.74)	NH 129.83 CO 176.13
Thr 73	7.53	4.59 (62.01)	4.66 (72.28)	C _{v2} H ₃ 1.55 (21.85) NH 112.55 CO 175.97
Asn 74	9.02	4.08 (55.79)	2.85, 2.85 (37.81)	N _{δ2} H 6.93, 7.59 (112.32) NH 119.25 CO 175.58
Leu 75	7.35	3.88 (56.95)	0.85, 1.20 (42.74)	C _v H 0.74 (26.32) C _{δ1} H ₃ 0.50 (25.19) C _{δ2} H ₃ 0.64 (23.97) NH 117.59 CO 176.62
Phe 76	6.81	4.53 (58.85)	3.01, 3.01 (43.03)	C _δ H 7.05, 7.05 C _{ε2} H 7.06, 7.06 NH 112.68 CO 176.19
His 77	9.02	4.33 (56.52)	3.25, 3.25 (27.52)	C _{ε1} H 8.61 (136.39) C _{δ2} H 7.19 (120.72) NH 114.65 CO 173.80
Cys 78	9.20	4.61 (56.80)	3.27, 3.49 (39.57)	NH 109.70 CO 175.34

Lys 79	7.20	4.37 (56.71)	2.26, 2.26 (32.76)	C _v H 1.58, 1.72 (24.74) C _o H 1.87, 2.00 (30.33) C _e H 3.31, 3.31 (41.95) NH 116.82 CO 176.30
Asp 80	8.56	4.20 (57.79)	2.80, 2.80 (40.42)	NH 121.71 CO 175.27
Lys 81	7.55	5.63 (54.83)	1.83, 1.94 (36.14)	C _v H 1.35, 1.35 (24.29) C _o H 1.64, 1.64 (29.63) C _e H 2.97, 2.97 (42.22) NH 114.43 CO 175.90
Asn 82	7.96	4.58 (54.96)	1.80, 2.10 (44.11)	N _{o2} H 3.64, 6.12 (115.84) NH 116.15 CO 173.41
Thr 83	7.81	4.57 (62.69)	2.48 (69.55)	C _v H ₃ 0.81 NH 122.40 CO 171.99
Phe 84	9.34	4.59 (56.30)	2.55, 2.98 (41.56)	C _{o1} H 6.95, 6.95 (132.40) C _{e1} H 6.94, 6.94 C _z H 6.44 (131.28) NH 125.70 CO 173.73
Ile 85	9.45	4.32 (60.83)	1.84 (40.66)	C _o H ₃ 0.42 (15.97) C _{v1} H 0.11, 1.14 (26.16) C _{v2} H 0.86 (18.95) NH 122.98 CO 175.67
Try 86	8.99	5.46 (55.28)	2.66, 3.20 (37.32)	C _e H 6.72, 6.72 (118.40) C _o H 6.98, 6.98 (133.00) NH 129.33 CO 173.42
Ser 87	7.73	4.39 (57.16)	3.23, 3.23 (65.17)	NH 121.04 CO 174.10
Arg 88	8.45	4.76 (56.21)	1.73, 2.44 (28.83)	C _v H 1.90, 1.90 (27.93) C _o H 3.34, 3.41 (43.03) NH 123.53
Pro 89		4.48 (66.15)	2.34, 2.50 (32.34)	C _v H 2.16, 2.33 (37.34) C _o H 3.95, 4.00 (50.37) CO 178.49
Glu 90	9.42	4.24 (63.59)	2.03, 2.36 (25.55)	C _v H 2.43, 2.67 (37.34) NH 118.42
Pro 91		4.49 (65.34)	1.93, 2.50 (31.64)	C _v H 2.11, 2.15 (28.09) C _o H 3.73, 3.99 (50.08) CO 178.38
Val 92	6.91	3.93 (64.98)	2.63 (31.81)	C _{v1} H ₃ 1.23 (23.12) C _{v2} H ₃ 1.28 (22.88) NH 115.94 CO 176.71
Lys 93	8.47	2.92 (59.24)	1.53, 1.62 (32.31)	C _v H -0.09, 1.17 (25.75) C _o H 1.18, 1.26 (29.16) C _e H 2.28, 2.43 (42.15) NH 121.42

				CO 179.62
Ala 94	7.74	3.93 (54.18)	1.48 (18.34)	NH 115.82 CO 172.44
Ile 95	7.42	3.71 (64.44)	1.77 (38.43)	C _δ H ₃ 0.80 (15.62) C _{γ1} H 1.28, 1.84 (27.46) C _{γ2} H 0.69 (17.05) NH 119.56 CO 177.24
Cys 96	7.15	4.72 (51.60)	1.86, 2.99 (36.30)	NH 111.73 CO 174.34
Lys 97	7.31	3.90 (59.35)	1.77, 1.84 (32.11)	C _γ H 1.39, 1.44 (24.48) C _δ H 1.66, 1.71 (29.55) C _ε H 2.98, 2.98 (42.18) NH 123.82 CO 177.76
Gly 98	9.16	3.52, 4.14 (45.29)		NH 114.01 CO 172.44
Ile 99	7.82	4.01 (60.94)	1.79 (36.33)	C _δ H ₃ 0.77 (12.41) C _{γ1} H 0.99, 1.28 (26.69) C _{γ2} H 0.78 (18.33) NH 122.34 CO 173.81
Ile 100	8.24	3.84 (63.40)	1.95 (37.78)	C _δ H ₃ 0.76 (11.70) C _{γ1} H 1.19, 1.41 (28.40) C _{γ2} H 0.97 (17.46) NH 128.11 CO 176.61
Ala 101	7.74	4.36 (52.03)	1.42 (18.66)	NH 123.50 CO 176.68
Ser 102	7.97	5.01 (58.68)	3.65, 3.65 (63.57)	NH 112.95 CO 175.74
Lys 103	8.34	4.37 (56.18)	1.58, 1.80 (36.03)	C _γ H 1.6, 1.36 (24.52) C _δ H 1.59, 1.65 (29.65) C _ε H 2.97, 2.97 (41.94) NH 124.74 CO 174.52
Asn 104	8.68	5.80 (52.15)	2.40, 2.77 (37.41)	N _{δ2} H 6.82, 7.50 (110.28) NH 125.81 CO 175.18
Val 105	9.41	4.15 (61.63)	1.81 (36.10)	C _{γ1} H ₃ 0.72 (21.36) C _{γ2} H ₃ 0.78 (21.81) NH 124.81 CO 172.59
Leu 106	8.12	4.65 (52.70)	1.42, 1.74 (43.86)	C _γ H 1.81 (26.04) C _{δ1} H ₃ 0.47 (23.17) C _{δ2} H ₃ 0.86 (26.35) NH 127.84 CO 177.20
Thr 107	8.67	4.10 (63.13)	4.65 (69.82)	C _{γ2} H ₃ 0.50 (21.80) NH 118.42 CO 174.67
Thr 108	8.51	4.24 (63.02)	4.37 (69.24)	C _{γ2} H ₃ 1.30 (22.19) NH 111.59

				CO 175.12
Ser 109	7.58	4.64 (56.65)	3.68, 3.68 (65.44)	NH 114.39 CO 172.75
Glu 110	8.40	4.18 (56.31)	1.67, 1.78 (30.73)	C _v H 2.13, 2.24 (36.76) NH 119.35 CO 175.64
Phe 111	8.58	4.63 (56.55)	3.03, 3.35 (44.11)	C _e H 7.23, 7.23 (131.41) C _o H 7.48, 7.48 (133.08) C _z H 7.18 (129.73) NH 121.25 CO 174.37
Tyr 112	9.71	4.94 (58.30)	3.03, 3.61 (38.12)	C _e H 6.96, 6.96 (118.70) C _o H 7.43, 7.43 (133.33) NH 119.47. CO 175.02
Leu 113	9.31	5.42 (53.87)	1.26, 2.00 (43.72)	C _v H 0.97 (27.57) C _{o1} H ₃ 0.75 (24.23) C _{o2} H ₃ 0.99 (24.00) NH 126.06 CO 178.10
Ser 114	9.49	5.21 (58.03)	3.73, 3.93 (66.30)	NH 121.01 CO 171.71
Asp 115	9.43	5.56 (53.24)	2.51, 2.62 (43.86)	NH 125.80 CO 176.00
Cys 116	8.47	5.30 (53.43)	2.46, 2.70 (39.54)	NH 118.90 CO 173.40
Asn 117	8.71	5.63 (52.12)	2.74, 2.88 (41.95)	N _{o2} H 7.06, 7.43 (111.22) NH 121.96 CO 174.69
Val 118	8.95	3.66 (62.65)	2.07 (31.94)	C _{v1} H ₃ 0.77 (21.61) C _{v2} H ₃ 0.88 (20.42) NH 127.57 CO 174.89
Thr 119	8.36	4.47 (60.94)	4.03 (70.37)	C _{v2} H ₃ 1.01 (20.68) NH 119.59 CO 173.09
Ser 120	7.75	4.26 (59.94)	3.83, 2.83 (64.99)	NH 122.94

Consejo Superior de Investigaciones Científicas
Instituto de Química Física “Rocasolano”
Departamento de Química Física Biológica
Grupo de Estructura, Dinámica e Interacciones de Proteínas por RMN



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

